ELSEVIER

Comput. Methods Appl. Mech. Engrg. 171 (1999) 175–204

Computer methods in applied mechanics and engineering

Adaptive numerical analysis in primal elastoplasticity with hardening

Jochen Alberty, Carsten Carstensen*, Darius Zarrabi

University of Kiel, Mathematisches Seminar II, Ludewig-Meyn-Str. 4, 24098 Kiel, Germany

Received 15 February 1998; revised 30 April 1998

Abstract

The quasi-static viscoplastic resp. elastoplastic evolution problem with isotropic or kinematic hardening is considered with emphasis on optimal convergence and adapted mesh-refining in the spatial discretization. Within one time-step of an implicit time-discretization, the finite element method leads to a minimisation problem for non-smooth convex functions on discrete subspaces. For piecewise constant resp. affine ansatz functions, the stress resp. displacement approximations are experimentally and theoretically shown to converge *linearly*. An a posteriori error estimate then justifies an automatic adaptive mesh-refining algorithm. Numerical examples support the superiority of the adapted mesh.

1. Introduction

In the engineering literature, the elastoplastic evolution problem is usually modelled with yield functions and flow rules written in terms of admissible stresses. In the discretization of this, Han and Reddy [1] call *dual* formulation, displacement and stress approximations are computed simultaneously. The elastoplastic material behaviour can equivalently be modelled by what Han and Reddy [1] call *primal* formulation where the strains are the primary variables and so a discretization requires the simultaneous approximation of the displacement and plastic strain field.

In this paper, we study the primal formulation and present a refined a priori and a posteriori error analysis for the spatial discretization. For a convenient reading, we outline both, the strain and stress formulation of plasticity, and their duality in Section 2. Within each time-step of the primal formulation in Section 3, we have to minimise a Lipschitz-continuous, but *non*-smooth convex functional over a linear space (not merely a convex set as in the dual method). The spatial Galerkin discretization in Section 4 is simple: replace the linear space by a conform FE-space.

Notice that the stress related dual model results in a minimisation of a quadratic problem over a convex set and so raises the question of the conform or non-conform approximation of sets. In particular, this is problematic for *hp*-methods (with higher order polynomial ansatz functions) where the discrete stress field satisfies the yield conditions in a finite set of discrete points rather than almost everywhere.

The error analysis of the time-discretized primal formulation, also called Hencky plasticity, time-independent plasticity, or holonomic plasticity, is the main topic of this paper. In the simplest spatial FE-discretization (piecewise affine displacements) with (largest) mesh-size h, the a priori error estimates in the literature (see e.g. [2] and the survey [1]) predict a convergence as $O(\sqrt{h})$. Since the functional to be minimised is non-smooth, we cannot expect a better convergence rate in general. However, in the dual stress formulation, the convergence

* Corresponding author.

0045-7825/99/$ – see front matter
PII: S0045-7825(98)00210-2

estimate is O(h) [3,4,23] which seemingly indicates superiority of the latter and more popular model. Since the two models are equivalent on the continuous level, their discrete counterparts are related and so we should expect a higher convergence rate of the discrete primal formulation.

The aim of this paper is to present theoretical and experimental evidence that the convergence rate in the primal formulation is, indeed, O(h). The proof is based on Jensen's inequality in Section 5. This result supports that the primal formulation is equally accurate as the dual formulation but favourable for higher order schemes. In a subsequent mathematical paper [5], we aim to prove this O(h) result also for the case of non-homogeneous data which, according to the subtle argument here, will be a laborious technical task.

Because of the limited regularity of the solution [6–9], adaptive mesh-refining appears to be a necessary tool for the efficient numerical treatment of elastoplastic problems [10,11]. In Section 6, we introduce an a posteriori error estimate which yields a computable upper error bound in the discrete primal formulation. The estimate is similar to the standard residual based a posteriori error indicator according to the hardening.

The numerical treatment of the discrete non-smooth problem is discussed in Section 7. We suggest a Newton–Raphson scheme which performs (sometimes) well in practise although we only offer little theoretical support for that. Numerical model examples in Section 8 confirm the improved convergence estimates and provide experimental evidence that the adaptive mesh-refining algorithm proposed leads to a more efficient computation.

The analysis of the fully numerical time-dependent problem with a complete refined error analysis will be the subject of a forthcoming paper [12].

2. Primal model of elastoplasticity

The strong formulation of small strain elastoplasticity with kinematic or isotropic hardening from the engineering literature is outlined and recast to the primal formulation under question here.

The elastoplastic body under consideration occupies a bounded Lipschitz domain Ω in $\mathbb{R}^d$. Local (quasi-static) equilibrium for the stress field $\sigma \in L^2(\Omega; \mathbb{R}^{d\times d})$ demands

$$\sigma = \sigma^{\mathrm{T}} \quad \text{and} \quad \operatorname{div} \sigma + f = 0 \quad \text{in } \Omega, \tag{2.1}$$

where f is the vector field of given body forces and we understand the divergence in (2.1) in the sense of distributions. The boundary Γ is split into a Dirichlet boundary Γ_{D}, a closed subset of Γ with positive surface measure, and the remaining (relatively open and possibly empty) Neumann part $\Gamma_{\mathrm{N}} := \Gamma \backslash \Gamma_{\mathrm{D}}$. We pose essential and static boundary conditions, namely,

$$u = 0 \quad \text{on } \Gamma_{\mathrm{D}} \quad \text{and} \quad \sigma \cdot n = g \quad \text{on } \Gamma_{\mathrm{N}}, \tag{2.2}$$

where g is a given applied surface force. With the displacement field $u \in H^1_{\mathrm{D}}(\Omega) := \{u \in H^1(\Omega)^d : u = 0 \text{ on } \Gamma_{\mathrm{D}}\}$ we associate the linear Green strain tensor

$$\varepsilon(u) := (\nabla u + (\nabla u)^{\mathrm{T}})/2 \quad \text{a.e. in } \Omega, \tag{2.3}$$

the symmetric part of the gradient of u. In the context of small strain elastoplasticity, the total strain $\varepsilon(u)$ is split additively into two contributions

$$\varepsilon(u) = \mathbb{C}^{-1}\sigma + p \quad \text{a.e. in } \Omega. \tag{2.4}$$

The elasticity operator is $\mathbb{C} : \mathbb{R}^{d\times d} \to \mathbb{R}^{d\times d}$,

$$\mathbb{C}\gamma = \lambda \operatorname{tr} \gamma I_d + 2\mu\gamma \ (\gamma \in \mathbb{R}^{d\times d}_{\mathrm{sym}}) \quad \text{for } \lambda, \mu > 0 \tag{2.5}$$

with the $d \times d$-unit matrix I_d and $\mathbb{C}^{-1}\sigma$ is the elastic and p is the plastic part of the total strain $\varepsilon(u)$. Note that the elastic material behaviour is characterised by $p = 0$ and that we need another material law to determine p. Moreover, there are restrictions on the stress variables prescribed by a dissipation functional φ which is convex and non-negative but may be $+\infty$. The first restriction is

$$\varphi(\sigma, \alpha) < \infty \quad \text{a.e. in } \Omega. \tag{2.6}$$

In this way, the hardening parameter $\alpha \in \mathbb{R}^m$ steers the set of admissible stresses; the couple (σ, α) is named *generalised stresses* and is called *admissible* if $\varphi(\sigma, \alpha) < \infty$.

The evolution of p and α is given by the Prandtl–Reuß normality law which states, for all other generalised stresses (τ, β), that there holds

$$\dot{p} : (\tau - \sigma) - \dot{\alpha} : (\beta - \alpha) \leqslant \varphi(\tau, \beta) - \varphi(\sigma, \alpha) \quad \text{a.e. in } \Omega . \tag{2.7}$$

Here, the dot denotes time derivatives, e.g. $\dot{p} := \partial p / \partial t$, and colon is a scalar product of matrices such that $A : B = \sum_{j,k=1}^{d} A_{jk} B_{jk}$ if $A, B \in \mathbb{R}^{d \times d}$; the Euclidean length $|\cdot|$ is defined by $|A| := \sqrt{A : A}$.

REMARK 2.1. According to the normality rule (2.7) the problem is time-dependent and so we are seeking variables (u, p, σ, α) in a time interval $[0, T]$ that satisfy consistent initial conditions at time $t = 0$ and (2.1)–(2.7) for almost all times in $(0, T)$. We refer to the mathematical literature [1,6–9] for details, existence and (non-) uniqueness, and for (poor) regularity properties of solutions.

REMARK 2.2. In the examples below, there is a given convex yield function Φ such that the admissible generalised stresses are characterised by

$$\Phi(\sigma, \alpha) \leqslant 0 \quad \text{in } \Omega \tag{2.8}$$

and φ is the characteristic functional of the set of admissible generalised stresses, i.e.

$$\varphi(\sigma, \alpha) := \begin{cases} 0 & \text{if } \Phi(\sigma, \alpha) \leqslant 0 , \\ \infty & \text{if } \Phi(\sigma, \alpha) > 0 . \end{cases} \tag{2.9}$$

In this case, (2.7) means that $\Phi(\sigma, \alpha) \leqslant 0$ and, for all (τ, β) with $\Phi(\tau, \beta) \leqslant 0$, there holds

$$\dot{p} : (\tau - \sigma) - \dot{\alpha} : (\beta - \alpha) \leqslant 0 \quad \text{a.e. in } \Omega . \tag{2.10}$$

An equivalent formulation, dual to (2.7), is obtained by using the dual φ^* of φ, defined by

$$\varphi^*(b) := \sup_a \{a \cdot b - \varphi(a)\} . \tag{2.11}$$

In terms of convex analysis, with ∂ denoting the subgradient, $a \in \partial\varphi(b)$ is equivalent to $b \in \partial\varphi^*(a)$. Therefore, the dual form to (2.7) reads

$$\sigma : (q - \dot{p}) + \alpha : (\beta + \dot{\alpha}) \leqslant \varphi^*(q, \beta) - \varphi^*(\dot{p}, -\dot{\alpha}) \quad \text{a.e. in } \Omega . \tag{2.12}$$

In the primal formulation under question here, we employ the normal rule (2.12) and require the dual φ^* of the dissipation functional.

DEFINITION 2.1 (Problem (P)). Seek (u, p, σ, α) satisfying consistent initial conditions and, in some time interval $(0, T)$, (2.1)–(2.4) and (2.12).

EXAMPLE 2.1 (Isotropic harding). Let $m = 1$, i.e. α is a scalar, and define

$$\Phi(\sigma, \alpha) := |\operatorname{dev} \sigma| - \sigma_y(1 + H\alpha) \tag{2.13}$$

in case $\alpha \geqslant 0$ (and $\Phi(\sigma, \alpha) = \infty$ if $\alpha < 0$ which, thereby, is not allowed). With the trace $\operatorname{tr} A := \sum_{j=1}^{d} A_{jj}$ and the $d \times d$-unit matrix I_d, the deviatoric part of a matrix is

$$\operatorname{dev} A := A - \frac{1}{d} (\operatorname{tr} A) I_d \quad (A \in \mathbb{R}^{d \times d}) .$$

The material constant $\sigma_y > 0$ is the yield stress and the constant $H > 0$ is the modulus of hardening. Then, there exists a unique solution of Problem (P) provided the exterior load f is slightly more regular (and then there holds Johnson's safe-load assumption) [3,4]. The dual functional is known (see e.g. [13] for a proof); for all $A \in \mathbb{R}^{d \times d}_{\mathrm{sym}}$ and $B \in \mathbb{R}$,

$$\varphi^*(A,B)=\begin{cases}\sigma_y|A| & \text{if } \operatorname{tr} A=0\wedge B+H\sigma_y|A|\leqslant 0\,,\\ \infty & \text{if } \operatorname{tr} A\neq 0\vee B+H\sigma_y|A|>0\,.\end{cases} \tag{2.14}$$

EXAMPLE 2.2 (Kinematic hardening). Let $m=d(d+1)/2$ and identify $\mathbb{R}^m\equiv\mathbb{R}^{d\times d}_{\mathrm{sym}}:=\{A\in\mathbb{R}^{d\times d}: A=A^{\mathrm{T}}\}$. Like the stress σ we consider α (pointwise) as a $\mathbb{R}^{d\times d}_{\mathrm{sym}}$-matrix and define

$$\Phi(\sigma,\alpha):=|\operatorname{dev}\sigma-\operatorname{dev}\alpha|-\sigma_y\,. \tag{2.15}$$

Then, there exists a unique solution of Problem (P) provided the exterior load f is slightly more regular (and then there holds Johnson's safe-load assumption) [1,3,4]. The dual functional equals (see e.g. [13] for a proof). for all $A,B\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$,

$$\varphi^*(A,B)=\begin{cases}\sigma_y|A| & \text{if } \operatorname{tr} A=0\wedge B=-A\,,\\ \infty & \text{if } \operatorname{tr} A\neq 0\vee B\neq -A\,.\end{cases} \tag{2.16}$$

EXAMPLE 2.3 (Perfect plasticity). In the case $m=0$ of no hardening, i.e. the internal variables are absent, the von-Mises yield condition is given by

$$\Phi(\sigma):=|\operatorname{dev}\sigma|-\sigma_y\,. \tag{2.17}$$

Although the resulting problem is covered in this section, the missing hardening leads to a different functional analytic frame. There exist solutions of Problem (P) in a much weaker sense (space of bounded deformation $BD(\Omega)$) if Johnson's safe-load assumption holds [8,9]. For any $A\in\mathbb{R}^{d\times d}_{\mathrm{sym}}$,

$$\varphi^*(A)=\begin{cases}\sigma_y|A| & \text{if } \operatorname{tr} A=0\,,\\ \infty & \text{if } \operatorname{tr} A\neq 0\,.\end{cases} \tag{2.18}$$

EXAMPLE 2.4 (Viscoplasticity). In Examples 2.1, 2.2 and 2.3 the dissipation functional (2.8) is non-smooth, but may be approximated by a smoother functional. The Yosida-regularisation leads to a viscoplastic material description in the sense of Perzyna where, given a viscosity $\mu>0$, for all preceding examples of Φ we define

$$\varphi(\sigma,\alpha):=\frac{1}{2\mu}\inf\{|(\sigma-\tau,\alpha-\beta)|^2:(\tau,\beta)\in\mathbb{R}^{d\times d}_{\mathrm{sym}}\times\mathbb{R}^m \text{ with } \Phi(\tau,\beta)\leqslant 0\}\,. \tag{2.19}$$

For $\mu>0$ there exists a unique solution of Problem (P) [8]. The dissipation functional (2.19) is, in some sense, converging towards (2.9) as $\mu\to 0$ [8]. Some calculations verify formulae for the dual functional, e.g. in perfect plasticity of Example 2.3, we obtain

$$\varphi^*(A)=\begin{cases}\sigma_y|A|+\dfrac{\mu}{2}|A|^2 & \text{if } \operatorname{tr} A=0\,,\\ \infty & \text{if } \operatorname{tr} A\neq 0\,.\end{cases} \tag{2.20}$$

3. Discretization in time

An implicit time-discretization (such as the implicit Euler method or the Crank–Nicholson scheme) yields in each time-step a spatial problem where we are given the variables $(u(t),\ p(t),\ \sigma(t),\ \alpha(t))$ at time $t=t_0$ as initial values, denoted as $(u_0,\ p_0,\ \sigma_0,\ \alpha_0)$, and seek corresponding approximations $(u_1,\ p_1,\ \sigma_1,\ \alpha_1)$ to $(u(t_1),\ p(t_1),\ \sigma(t_1),\ \alpha(t_1))$ at time $t_1=t_0+k$. Therein, time derivatives such as $\dot p(t_1)$ are replaced by backward difference quotients as $(p_1-p_0)/k$, $k:=t_1-t_0>0$. Rewriting (2.1)–(2.2) with $\sigma=\mathbb{C}[\varepsilon(u)-p]$ in the standard weak form, known as the principle of virtual work in mechanics, we deduce the following time-discretized problem.

DEFINITION 3.1 (Variational Problem (P_1)). Seek $(u_1,\ p_1,\ \alpha_1)\in H^1_{\mathrm{D}}(\Omega)\times L^2(\Omega)^{d\times d}_{\mathrm{sym}}\times L^2(\Omega)^m$ satisfying, for all $(v,\ q,\ \beta)\in H^1_{\mathrm{D}}(\Omega)\times L(\Omega)^{d\times d}_{\mathrm{sym}}\times L(\Omega)^m$,

$$\int_\Omega \mathbb{C}[\varepsilon(u_1)-p_1]:\varepsilon(v)\,\mathrm{d}x=\int_\Omega fv\,\mathrm{d}x+\int_{\Gamma_{\mathrm{N}}} gv\,\mathrm{d}s\,, \tag{3.1}$$

$$\int_\Omega \{\mathbb{C}[\varepsilon(u_1) - p_1] : (kq - p_1 + p_0) + \alpha_1 : (k\beta + \alpha_1 - \alpha_0)\}\,dx$$
$$\leq k\int_\Omega \varphi^*(q, \beta)\,dx - k\int_\Omega \varphi^*((p_1 - p_0)/k, (\alpha_0 - \alpha_1)/k)\,dx\,. \tag{3.2}$$

The variational problem (P_1) is equivalent to the following minimisation problem.

DEFINITION 3.2 (Minimisation Problem (M_1)). Seek a minimiser $(u_1, p_1, \alpha_1) \in H^1_D(\Omega) \times L^2(\Omega)^{d\times d}_{sym} \times L^2(\Omega)^m$ of

$$\frac{1}{2}\int_\Omega \mathbb{C}[\varepsilon(u_1) - p_1 + p_0] : (\varepsilon(u_1) - p_1 + p_0)\,dx + \frac{1}{2}\int_\Omega |\alpha_1|^2\,dx$$
$$+ k\int_\Omega \varphi^*\Big(\frac{p_1 - p_0}{k}, \frac{\alpha_0 - \alpha_1}{k}\Big)\,dx - \int_\Omega f u_1\,dx - \int_{\Gamma_N} g u_1\,ds - \int_\Omega p_0 : \mathbb{C}[\varepsilon(u_1) - p_1]\,dx\,. \tag{3.3}$$

REMARK 3.1. The problem (3.3) is local in α_1 in the sense that it can be solved separately for each material point [13]. Indeed, for each $p := p_1 - p_0 \in \mathbb{R}^{d\times d}_{sym}$, direct calculations show that the condition for α_1 in (3.2) resp. (3.3) is equivalent to the minimisation

$$\min_{a\in\mathbb{R}^m}\left\{\frac{1}{2}|a|^2 + k\varphi^*(p/k, (\alpha_0 - a)/k)\right\} = j(x, p) + \frac{\eta}{2}|p|^2 + \frac{\nu}{2}|\alpha_0|^2\,, \tag{3.4}$$

where the minimum is attained for $a = \alpha_1$. Here, $\eta \geq 0$ is some hardening constant and $j(x, \cdot)$ is a convex functional which depends on x through $\alpha_0(x)$. In the examples of this paper, $j(x, \cdot)$ has the form

$$j(x, p) = \begin{cases} m(x)|p| + M(x) : p & \text{if tr } p = 0\,, \\ \infty & \text{if tr } p \neq 0\,, \end{cases} \quad (x \in \Omega;\ p \in \mathbb{R}^{d\times d}_{sym})\,, \tag{3.5}$$

with functions m and M shown in Table 1. (For isotropic hardening, we suppose $\alpha_0 \geq 0$.)

A change of variables $p := p_1 - p_0$, $u := u_1$, and employing (3.4), we obtain an equivalent minimisation problem [13].

DEFINITION 3.3 (Minimisation Problem (M_2)). Seek a minimiser (u, p) of

$$f(u, p) := \frac{1}{2}\int_\Omega \mathbb{C}[\varepsilon(u) - p] : (\varepsilon(u) - p)\,dx + \frac{1}{2}\int_\Omega \eta|p|^2\,dx + \int_\Omega j(p)\,dx - l(u, p) \tag{3.6}$$

in $H^1_D(\Omega) \times L^2(\Omega)^{d\times d}_{sym}$, where the linear functional l is defined by

$$l(u, p) := \int_\Omega (fu + p_0 : \mathbb{C}[\varepsilon(u) - p])\,dx + \int_{\Gamma_N} gu\,ds\,. \tag{3.7}$$

A variation of u resp. p in the problem (M_2) leads to an equivalent variational inequality similar to (P_1).

Table 1
Parameter m and M in Eq. (3.5) and η, ν in Eq. (3.4)

$j(x, p) = m(x)|p| + M(x) : p + \frac{\nu}{2}|\alpha_0|^2$

	m	M	η	ν
Isotropic hardening	$(1 + H\alpha_0)\sigma_y$	0	$H^2\sigma_y^2$	1
Kinematic hardening	σ_y	α_0	1	1
Perfect plasticity	σ_y	0	0	0
Viscoplasticity	σ_y	0	μ/k	0

DEFINITION 3.4 (Variational Problem (P_2)*).* Seek $(u, p) \in H^1_D(\Omega) \times L^2(\Omega)^{d\times d}_{sym}$ satisfying, for all $(v, q) \in H^1_D(\Omega) \times L^2(\Omega)^{d\times d}_{sym}$,

$$\int_\Omega \mathbb{C}[\varepsilon(u) - p] : \varepsilon(v)\, dx = l(v, 0)\,, \tag{3.8}$$

$$\int_\Omega \{\mathbb{C}\varepsilon(u) - (\mathbb{C} + \eta)p\} : (q - p)\, dx + l(0, q - p) \leqslant \int_\Omega \{j(q) - j(p)\}\, dx\,. \tag{3.9}$$

4. Discretization in space

The discretization of Problem (P_2) consists of replacing $H^1_D(\Omega) \times L^2(\Omega)^{d\times d}_{sym}$ by a discrete subspace $\mathscr{H} \times \mathscr{L}$. In the simplest example we study conform piecewise affine resp. constant finite element test and trial functions.

Let $\mathscr{T}$ be a regular triangulation in triangles, tetrahedrons, parallelograms or quadrilaterals in the sense of Ciarlet [14,15], i.e. if $d = 2$, $\mathscr{T}$ is a finite partition of Ω in closed triangles or parallelograms; two distinct elements T_1 and T_2 in $\mathscr{T}$ are either disjoint, or $T_1 \cap T_2$ is a complete edge or a common node of both T_1 and T_2. With $\mathscr{T}$ let $\mathscr{E}$ denote the set of all edges, and we assume that $E \in \mathscr{E}$ either belongs to Γ_D or the surface measure of $E \cap \Gamma_D$ vanishes, so there is no change of boundary conditions within one edge $E \subseteq \Gamma$. Define by $\mathscr{S}^0(\mathscr{T})$ the $\mathscr{T}$-piecewise constant and by $\mathscr{S}^1(\mathscr{T})$ the continuous and $\mathscr{T}$-piecewise affine functions. Then, let

$$\mathscr{H} := \mathscr{S}^1(\mathscr{T})^d \cap H^1_D(\Omega) \quad \text{and} \quad \mathscr{L} := \operatorname{dev}(\mathscr{S}^0(\mathscr{T})^{d\times d}_{sym}) \tag{4.1}$$

and notice that nodal bases of $\mathscr{H}$ and $\mathscr{L}$ can be implemented very easily.

DEFINITION 4.1 (Discrete Problem $(P_2 | \mathscr{H} \times \mathscr{L})$*).* Seek $(U, P) \in \mathscr{H} \times \mathscr{L}$ satisfying, for all $(V, Q) \in \mathscr{H} \times \mathscr{L}$,

$$\int_\Omega \mathbb{C}[\varepsilon(U) - P] : \varepsilon(V)\, dx = l(V, 0)\,, \tag{4.2}$$

$$\int_\Omega \{\mathbb{C}\varepsilon(U) - (\mathbb{C} + \eta)P\} : (Q - P)\, dx + l(0, Q - P) \leqslant \int_\Omega \{j(Q) - j(P)\}\, dx\,. \tag{4.3}$$

REMARK 4.1. The discrete problem $(P_2 \,|\, \mathscr{H} \times \mathscr{L})$ is equivalent to a minimisation of (3.6) in $\mathscr{H} \times \mathscr{L}$ and therefore has unique solutions if $\eta > 0$.

5. A priori error analysis

Suppose that $(u, p) \in H^1_D(\Omega) \times L^2(\Omega)^{d\times d}_{sym}$ solves Problem (P_2) and that $(U, P) \in \mathscr{H} \times \mathscr{L}$ solves the discrete Problem $(P_2 \,|\, \mathscr{H} \times \mathscr{L})$. Let

$$\sigma := \mathbb{C}[\varepsilon(u) - p - p_0] \quad \text{resp.} \quad \Sigma := \mathbb{C}[\varepsilon(U) - P - \overline{p_0}]$$

be the exact resp. discrete stress field. Here, $\overline{p_0} \in \mathscr{S}^0(\mathscr{T})$ is defined by $\overline{p_0}|_T := \int_T p_0\, dx/\operatorname{meas}(T)$. Notice that

$$\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|^2_{L^2(\Omega)} = \int_\Omega \mathbb{C}^{-1}[\sigma - \Sigma] : (\sigma - \Sigma)\, dx \tag{5.1}$$

defines the energy norm of the stress error.

THEOREM 5.1. Suppose that j *is given as in (3.5). We have*

$$\begin{aligned}
&\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|^2_{L^2(\Omega)} + 2\|\eta^{1/2}[p - P]\|^2_{L^2(\Omega)} \\
&\quad \leqslant 2 \inf_{W\in\mathscr{H}} \|\mathbb{C}^{1/2}\varepsilon(u - W)\|^2_{L^2(\Omega)} + \inf_{Q\in\mathscr{L}} \|p - Q\|^2_{L^2(\Omega)} \\
&\qquad + 2 \inf_{K\in\mathscr{L}} \|M - K\|^2_{L^2(\Omega)} + 2 \inf_{K\in\mathscr{S}^0(\mathscr{T})} \|m - K\|^2_{L^2(\Omega)} + 4\mu \inf_{K\in\mathscr{L}} \|\mathbb{C}^{1/2}[p_0 - K]\|^2_{L^2(\Omega)}\,.
\end{aligned} \tag{5.2}$$

PROOF. The proof relies on Jensen's inequality for the convex functional $j(x,\cdot)$: The integral mean $\overline{p}|_T := \int_T p(y)\,dy/\mathrm{meas}(T)$ of p over the domain $T \in \mathcal{T}$ yields a dissipation $j(x,\overline{p})$ which is smaller than the integral mean of the dissipation $\int_T j(x, p(y))\,dy/\mathrm{meas}(T)$. The resulting elementwise inequality

$$\int_T j(x,\overline{p})\,dy \leq \int_T j(x, p(y))\,dy \tag{5.3}$$

(where $\overline{p} \in \mathcal{S}^0(\mathcal{T})$ is constant on T) is caused by the convexity of $j(x,\cdot)$ for each parameter $x \in \Omega$. Indeed, $\overline{p}|_T$ is a (limit of) convex combinations $\Sigma_j \lambda_j p_j$ which, owing to the concept of convexity, fulfil $j(x, \Sigma_k \lambda_k p_k) \leq \Sigma_k \lambda_k j(x, p_k)$. These inequalities then prove (5.3) for simple functions if we interpret λ_j as $\mathrm{meas}(\{y \in T \mid p(y) = p_j\})/\mathrm{meas}(T)$. A passage to the limit proves (5.3) for all L^2-functions. We refer to standard literature in mathematical analysis for further details.

Integrating (5.3) over $x \in T$ we infer

$$\int_T \{j(x,\overline{p}) - j(x, p(x))\}\,dx \leq \int_T \int_T \{j(x, p(y)) - j(x, p(x))\}\,dx\,dy/\mathrm{meas}(T)\,. \tag{5.4}$$

Assuming (3.5) and writing $q(x) := |p(x)|$ with integral mean $\overline{q} := \int_T q(x)\,dx/\mathrm{meas}(T)$, we recast the right-hand side in (5.4) as

$$\begin{aligned}&\int_T \int_T m(x)\{q(y) - q(x)\}\,dx\,dy/\mathrm{meas}(T) + \int_T \int_T M(x) : \{p(y) - p(x)\}\,dx\,dy/\mathrm{meas}(T)\\ &\quad = \int_T m(x)\{\overline{q} - q(x)\}\,dx + \int_T M(x) : \{\overline{p} - p(x)\}\,dx\,.\end{aligned} \tag{5.5}$$

Writing $\overline{m}$ resp. $\overline{M}$ for the integral means of m resp. M on T we make use of $\int_T \overline{M} : \{\overline{p} - p(x)\}\,dx = 0$, etc. and so deduce that (5.5) equals

$$\begin{aligned}&\int_T \{m(x) - \overline{m}\}\{\overline{q} - q(x)\}\,dx + \int_T \{M(x) - \overline{M}\} : \{\overline{p} - p(x)\}\,dx\\ &\quad \leq \|m - \overline{m}\|_{L^2(\Omega)}\|q - \overline{q}\|_{L^2(\Omega)} + \|M - \overline{M}\|_{L^2(\Omega)}\|p - \overline{p}\|_{L^2(\Omega)}\,.\end{aligned} \tag{5.6}$$

According to the triangle inequality, we conclude $\|q - \overline{q}\|_{L^2(\Omega)} \leq \|p - \overline{p}\|_{L^2(\Omega)}$ and so finally obtain

$$\int_T \{j(x,\overline{p}) - j(x, p(x))\}\,dx \leq \|p - \overline{p}\|_{L^2(\Omega)}(\|M - \overline{M}\|_{L^2(\Omega)} + \|m - \overline{m}\|_{L^2(\Omega)})\,. \tag{5.7}$$

The remaining arguments are standard: A subtraction of (3.8) by (4.2) yields the Galerkin orthogonality

$$\int_\Omega (\sigma - \Sigma) : \varepsilon(U - W)\,dx = 0 \tag{5.8}$$

for all $W \in \mathcal{H}$. Further, (3.9) and (4.3) lead to

$$\int_\Omega (\sigma - \eta p) : (P - p)\,dx \leq \int_\Omega \{j(x, P) - j(x, p)\}\,dx\,, \tag{5.9}$$

$$\int_\Omega (\Sigma - \eta P) : (\overline{p} - P)\,dx \leq \int_\Omega \{j(x,\overline{p}) - j(x, P)\}\,dx\,. \tag{5.10}$$

Utilising (5.7)–(5.10) in the identity

$$\begin{aligned}\|\mathbb{C}^{-1/2}[\sigma - \Sigma]\|^2_{L^2(\Omega)} &= \int_\Omega (\sigma - \Sigma) : \varepsilon(u - U)\,dx\\ &\quad - \int_\Omega \sigma : (p_0 - \overline{p_0})\,dx + \int_\Omega \sigma : (P - p)\,dx + \int_\Omega \Sigma : (\overline{p} - P)\,dx\end{aligned} \tag{5.11}$$

(because $\int_T \Sigma : \overline{p_0}\,dx = \int_T \Sigma : p_0\,dx$ for $\Sigma \in \mathcal{S}^0(\mathcal{T})$) we deduce

$$\|\mathbb{C}^{-1/2}[\sigma - \Sigma]\|^2_{L^2(\Omega)} + \|\eta^{1/2}[p - P]\|^2_{L^2(\Omega)} \leq \int_\Omega (\sigma - \Sigma) : \varepsilon(u - W)\,dx - \int_\Omega \sigma : (p_0 - \overline{p_0})\,dx + \|p - \overline{p}\|_{L^2(\Omega)}(\|M - \overline{M}\|_{L^2(\Omega)} + \|m - \overline{m}\|_{L^2(\Omega)}). \tag{5.12}$$

Writing $\overline{\sigma}$ resp. $\overline{p}$ for the elementwise integral mean of σ resp. p, we have

$$-\int_\Omega \sigma : (p_0 - \overline{p_0})\,dx = -\int_\Omega (\sigma - \overline{\sigma}) : (p_0 - \overline{p_0})\,dx \leq \|\mathbb{C}^{1/2}(p_0 - \overline{p_0})\|_{L^2(\Omega)}\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|_{L^2(\Omega)}. \tag{5.13}$$

With (5.13) and Cauchy's inequality, we infer that the right-hand side of (5.12) is upper bounded by

$$\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|_{L^2(\Omega)}(\|\mathbb{C}^{1/2}\varepsilon(u - W)\|_{L^2(\Omega)} + \|\mathbb{C}^{1/2}(p_0 - \overline{p_0})\|_{L^2(\Omega)}) + \|p - \overline{p}\|_{L^2(\Omega)}(\|M - \overline{M}\|_{L^2(\Omega)} + \|m - \overline{m}\|_{L^2(\Omega)}).$$

(Notice that, since $\operatorname{tr} p_0 = 0 = \operatorname{tr}\overline{p_0}$, $\mathbb{C}(p_0 - \overline{p_0}) = 2\mu(p_0 - \overline{p_0})$.) With Young's inequality, $ab \leq (a^2 + b^2)/2$, and subtracting $\|\mathbb{C}^{-1/2}[\sigma - \Sigma]\|^2_{L^2(\Omega)}/2$ we finally conclude (5.2). □

REMARK 5.1. The a priori error estimate (5.2) is quasi-optimal and shows linear convergence if u and p as well as the data are smooth.

REMARK 5.2. Notice that the constants in front of the right-hand side in (5.2) are independent of the hardening and so (5.2) holds for perfect plasticity as well. In the latter case $\eta = 0$, the upper bound in (5.2) may be infinity if u is only in $BD(\Omega)$ rather than in $H^1(\Omega)$.

REMARK 5.3. For Hencky plasticity or if the initial values are given on a coarser mesh (if the mesh is only allowed to be refined—no coarsening), p_0 and α_0 are constant on each element. Then, the preceding analysis results in

$$\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|^2_{L^2(\Omega)} + 2\|\eta^{1/2}(p - P)\|^2_{L^2(\Omega)} \leq \inf_{W \in \mathcal{W}} \|\mathbb{C}^{1/2}\varepsilon(u - W)\|^2_{L^2(\Omega)}. \tag{5.14}$$

REMARK 5.4. At first glance, it seems to be a surprising consequence of (5.14) that the approximation error of the plastic strain p does not affect the stress error. Notice that we used a mesh for the displacement field which is coarser than (in our analysis indeed equal to) the mesh for the plastic strain approximations. From (5.14) we observe that any h-, p, or hp-refinement of the mesh for the plastic strains is pointless. This overrules a consequence of the prior analysis [1,2]: the strain approximations need not to be further improved.

REMARK 5.5. From (5.2), we easily derive error estimates for the total strain field which relies on the hardening. Indeed, from the triangle inequality and $\mathbb{C}^{1/2}\varepsilon(u - U) = \mathbb{C}^{-1/2}(\sigma - \Sigma) + \mathbb{C}^{1/2}(p - P) + \mathbb{C}^{1/2}(p_0 - \overline{p_0})$ we infer

$$\|\mathbb{C}^{1/2}\varepsilon(u - U)\|^2_{L^2(\Omega)} \leq 2(1 + \mu/\eta)(\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|^2_{L^2(\Omega)} + \|\eta^{1/2}(p - P)\|^2_{L^2(\Omega)} + \|\mathbb{C}^{1/2}(p_0 - \overline{p_0})\|^2_{L^2(\Omega)}). \tag{5.15}$$

REMARK 5.6. For an a priori analysis of the time-discretization error, we refer to [1,12].

6. A posteriori error analysis and adaptive mesh-refining

For each $T \in \mathcal{T}$, let h_T denote its diameter and define

$$\eta_T^2 = h_T^2 \int_T |f + \operatorname{div}_{\mathcal{T}}\Sigma|^2\,dx + \int_{\partial T} h_E |J(\Sigma \cdot n_E)|^2\,ds + \inf_{Q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}} \|p_0 - Q\|^2_{L^2(T)} + \inf_{Q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}} \|M - Q\|^2_{L^2(T)} + \inf_{N \in \mathbb{R}} \|m - N\|^2_{L^2(T)}. \tag{6.1}$$

Here, $J(\Sigma \cdot n_E)$ is the jump of the discrete stress field Σ along an edge $E \in \mathscr{E}$ with normal n_E and size h_E with the usual modification $J(\Sigma \cdot n_E) := \Sigma \cdot n_E - g$ if $E \subset \Gamma_N$.

THEOREM 6.1. There exists an η-depending constant $C(\eta) > 0$ such that

$$\|\sigma - \Sigma\|^2_{L^2(\Omega)} + \|p - P\|^2_{L^2(\Omega)} + \|\varepsilon(u - U)\|^2_{L^2(\Omega)} \leq C(\eta) \sum_{T \in \mathscr{T}} \eta_T^2 . \tag{6.2}$$

PROOF. From (5.12) and (5.13) we have, for all $W \in \mathscr{H}$,

$$\|\mathbb{C}^{-1/2}[\sigma - \Sigma]\|^2_{L^2(\Omega)} + \|\eta^{1/2}[p - P]\|^2_{L^2(\Omega)} \leq \int_\Omega (\sigma - \Sigma) : \varepsilon(u - W)\,dx - \int_\Omega (\sigma - \bar{\sigma}) : (p_0 - \bar{p_0})\,dx + \|p - \bar{p}\|_{L^2(\Omega)}(\|M - \bar{M}\|_{L^2(\Omega)} + \|m - \bar{m}\|_{L^2(\Omega)}) . \tag{6.3}$$

Since $\sigma - \Sigma$ is symmetric we have $(\sigma - \Sigma) : \varepsilon(u - W) = (\sigma - \Sigma) : \nabla(u - W)$, and an elementwise integration by parts shows

$$\int_\Omega (\sigma - \Sigma) : \varepsilon(u - W)\,dx = \int_\Omega R(u - W)\,dx - \int_{\cup \mathscr{E}} J(u - W)\,ds , \tag{6.4}$$

where $\cup \mathscr{E}$ is the skeleton of all edges in $\mathscr{T}$. The volume residual is

$$R := f + \operatorname{div}_{\mathscr{T}} \Sigma \in L^2(\Omega) , \tag{6.5}$$

where $\operatorname{div}_{\mathscr{T}}$ denotes the elementwise divergence. Recall that the jump residual $J \in L^2(\cup \mathscr{E})$ is defined by

$$J|_E := \begin{cases} [\Sigma \cdot n_E] & \text{if } E \not\subseteq \Gamma , \\ 0 & \text{if } E \subseteq \overline{\Gamma_D} , \\ g - \Sigma \cdot n & \text{if } E \subseteq \overline{\Gamma_N} , \end{cases} \quad (E \in \mathscr{E}) . \tag{6.6}$$

For some weak interpolation W to u, the error $u - W$ obeys estimates of the form

$$\|h_{\mathscr{T}}^{-1}(u - W)\|^2_{L^2(\Omega)} + \|h_{\mathscr{E}}^{-1/2}(u - W)\|^2_{L^2(\cup \mathscr{E})} \leq c_1^2 \|\nabla(u - U)\|^2_{L^2(\Omega)} . \tag{6.7}$$

Here, $h_{\mathscr{T}}$ and $h_{\mathscr{E}}$ are piecewise constant, $h_{\mathscr{T}}|_T = h_T$ and $h_{\mathscr{E}}|_E = h_E$. We refer to [16–19] for details. Employing (6.7) in (6.4) shows, with Cauchy's inequality,

$$\int_\Omega (\sigma - \Sigma) : \varepsilon(u - W)\,dx \leq \sqrt{2} c_1 \|\nabla(u - U)\|_{L^2(\Omega)} (\|h_{\mathscr{T}} R\|^2_{L^2(\Omega)} + \|h_{\mathscr{E}}^{1/2} J\|^2_{L^2(\cup \mathscr{E})})^{1/2} . \tag{6.8}$$

Since Γ_D has a positive surface measure, Korn's inequality shows

$$\|\nabla(u - U)\|_{L^2(\Omega)} \leq c_2 \|\mathbb{C}^{1/2} \varepsilon(u - U)\|_{L^2(\Omega)} . \tag{6.9}$$

Hence, from (6.3), (6.8) and (6.9), we have

$$\begin{aligned} &\|\mathbb{C}^{-1/2}[\sigma - \Sigma]\|^2_{L^2(\Omega)} + \|\eta^{1/2}[p - P]\|^2_{L^2(\Omega)} \\ &\quad \leq \sqrt{2} c_1 c_2 \|\mathbb{C}^{1/2} \varepsilon(u - U)\|_{L^2(\Omega)} (\|h_{\mathscr{T}} R\|^2_{L^2(\Omega)} + \|h_{\mathscr{E}}^{1/2} J\|^2_{L^2(\cup \mathscr{E})})^{1/2} \\ &\qquad + \|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|_{L^2(\Omega)} \|\mathbb{C}^{1/2}(p_0 - \bar{p_0})\|_{L^2(\Omega)} \\ &\qquad + \|p - \bar{p}\|_{L^2(\Omega)} (\|M - \bar{M}\|_{L^2(\Omega)} + \|m - \bar{m}\|_{L^2(\Omega)}) . \end{aligned} \tag{6.10}$$

With Young's inequality and some η-depending constants $c_1(\eta)$, $c_2(\eta)$, and $c_3(\eta)$ we deduce from (5.15)

$$\begin{aligned} &\|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|^2_{L^2(\Omega)} + \|\eta^{1/2}(p - P)\|^2_{L^2(\Omega)} + \|\mathbb{C}^{1/2} \varepsilon(u - U)\|^2_{L^2(\Omega)} \\ &\quad \leq \frac{1}{2} \|\mathbb{C}^{1/2} \varepsilon(u - U)\|^2_{L^2(\Omega)} + c_1(\eta)^2 (\|h_{\mathscr{T}} R\|^2_{L^2(\Omega)} + \|h_{\mathscr{E}}^{1/2} J\|^2_{L^2(\cup \mathscr{E})}) \\ &\qquad + \frac{1}{2} \|\mathbb{C}^{-1/2}(\sigma - \Sigma)\|^2_{L^2(\Omega)} + c_2(\eta) \|\mathbb{C}^{1/2}(p_0 - \bar{p_0})\|^2_{L^2(\Omega)} + \frac{1}{2} \|\eta^{1/2}(p - P)\|^2_{L^2(\Omega)} \\ &\qquad + c_3(\eta)(\|M - \bar{M}\|_{L^2(\Omega)} + \|m - \bar{m}\|_{L^2(\Omega)})^2 . \quad \square \end{aligned} \tag{6.11}$$

REMARK 6.1. If p_0 and α_0 are constant on T, for instance in time-independent problems ($p_0 = 0 = \alpha_0$) or if the mesh is refined successively ($p_0, \alpha_0 \in \mathscr{P}^0(\mathscr{T})$), the error indicator η_T reduces to

$$\eta_T^2 = h_T^2 \int_T |f + \operatorname{div}_{\mathscr{T}} \Sigma|^2 \, dx + \int_{\partial T} h_E |J(\Sigma \cdot n_E)|^2 \, ds \tag{6.12}$$

and so is the same as in pure elasticity (utilising the stress field from a discrete elastoplastic problem).

REMARK 6.2. In perfect plasticity, Theorem 6.1 is expected to be false. A closer inspection then shows that it is required to follow the arguments in [11] and derive weaker estimates. Essentially it is because (5.15) is not available and so $\|\nabla(u - U)\|_{L^2(\Omega)}$ cannot be absorbed.

REMARK 6.3. Employing the arguments of [16,17], we could sharpen the estimate of Theorem 6.1 which shows that, generically, the volume contribution $h_T^2 \int_T |f + \operatorname{div}_{\mathscr{T}} \Sigma|^2 \, dx$ can be neglected (replaced by a higher-order term).

REMARK 6.4. The reliable inequality is sharp in the sense that there holds the reverse inequality up to data approximations, i.e.

$$\begin{aligned} \eta_T^2 \leqslant C(\eta)(&\|\sigma - \Sigma\|_{L^2(\omega)}^2 + \|p - P\|_{L^2(\omega)}^2 + \|\varepsilon(u - U)\|_{L^2(\omega)}^2 \\ &+ \inf_{Q \in \mathbb{R}^{d \times d}_{\mathrm{sym}}} \|p_0 - Q\|_{L^2(\omega)}^2 + \inf_{Q \in \mathbb{R}^{d \times d}_{\mathrm{sym}}} \|M - Q\|_{L^2(\omega)}^2 + \inf_{N \in \mathbb{R}} \|m - N\|_{L^2(\omega)}^2) \end{aligned} \tag{6.13}$$

for each element $T \in \mathscr{T}$ and its neighbourhood $\omega := \cup \{T' \in \mathscr{T} : T \cap T' \neq \emptyset\}$ and for some η-depending constant $C(\eta) > 0$. The proof is as in [19] and we refer to [5] for details.

ALGORITHM 6.1.

(a) Start with a coarse mesh $\mathscr{T}_0$, $k = 0$.
(b) Solve the discrete problem with respect to the actual mesh $\mathscr{T}_k$.
(c) Compute η_T for all $T \in \mathscr{T}_k$.
(d) Compute a given stopping criterion and decide to terminate or to continue and then go to (e).
(e) Mark the element T for (*red*) refinement provided

$$\eta_T \geqslant \frac{1}{2} \max_{T' \in \mathscr{T}_k} \eta_{T'} .$$

(f) Mark further elements (within a *red–green–blue* refinement) to avoid hanging nodes. Define the resulting mesh as the actual mesh $\mathscr{T}_{k+1}$, update k and go to (b).

Details on the so-called *red–green–blue* refinement strategies may be found in [19].

REMARK 6.5. In the numerical examples below, the volume contributions to the error indicator η_T in (6.12) vanish according to $f = 0$, $\bar{p}_0|_T$ is constant and $U|_T$ affine. Hence, we may argue as in [19] and obtain equivalence of the edge contributions to the ZZ-estimator. In this way, we justify the popular ZZ-estimator in viscoplasticity and plasticity with hardening (in cases where the influence of $\|p - \bar{p}_0\|_{L^2(T)}$ is negligible). According to Remark 6.3, this argument is valid also in case that f is non-zero but smooth. The authors are unaware of any other justification of the reliability of the ZZ-estimator in the context of plasticity. However, the constants are η-depending and so the reliability of the ZZ-estimator in perfect plasticity remains as an open question.

7. Numerical solution algorithms

The numerical treatment of $(P_2 \mid \mathscr{H} \times \mathscr{L})$ is simplified by the elimination of the variable P. Indeed, given $\varepsilon(U)$ we can solve $(P_2 \mid \mathscr{H} \times \mathscr{L})$ elementwise.

PROPOSITION 7.1. Given $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *and* $b > 0$ *there exists exactly one* $P \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *with* $\operatorname{tr} P = 0$ *that satisfies*

$$\{A - (\mathbb{C} + \eta)P\} : (Q - P) \leqslant b\{|Q| - |P|\} \tag{7.1}$$

for all $Q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ *with* $\operatorname{tr} Q = 0$. *This* P *is characterised as the minimiser of*

$$\frac{1}{2}(\mathbb{C} + \eta)P : P - P : A + b|P| \tag{7.2}$$

(amongst trace-free symmetric $d \times d$*-matrices) and equals*

$$\frac{(|\operatorname{dev} A| - b)_+}{2\mu + \eta} \frac{\operatorname{dev} A}{|\operatorname{dev} A|}, \tag{7.3}$$

where $(\cdot)_+ := \max\{0, \cdot\}$ *denotes the non-negative part. The minimal value of (7.2) (attained for* P *as in (7.3)) is*

$$-\frac{1}{2}(|\operatorname{dev} A| - b)_+^2/(2\mu + \eta). \tag{7.4}$$

PROOF. In convex analysis, (7.1) states that

$$A - (\mathbb{C} + \eta)P \in b\partial|\cdot|(P) \tag{7.5}$$

where $\partial|\cdot| = \operatorname{sign}$ denotes the subgradient of the modulus function, and only trace-free arguments are under consideration. The modulus function is convex and so is (7.2). Identity (7.5) is equivalent to 0 belonging to the subgradient of (7.2), which characterises the minimisers of (7.2). If $P = 0$, (7.1) states

$$A : Q \leqslant b|Q| \tag{7.6}$$

for all $Q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ with $\operatorname{tr} Q = 0$. Hence, $|\operatorname{dev} A| \leqslant b$. If $|\operatorname{dev} A| > b$ we conclude $P \neq 0$ and obtain $\partial|\cdot|(P) = \{P/|P|\}$. Hence, (7.5) yields

$$\operatorname{dev} A - (\mathbb{C} + \eta)P = bP/|P|. \tag{7.7}$$

Notice that $\operatorname{tr} \mathbb{C}P = 0$ as $\operatorname{tr} P = 0$, and only trace-free arguments are under consideration. Since then $\mathbb{C}P = 2\mu P$ we obtain

$$\operatorname{dev} A = (b + (2\mu + \eta)|P|)P/|P| \tag{7.8}$$

and so $|\operatorname{dev} A| = b + (2\mu + \eta)|P|$, whence

$$|P| = (|\operatorname{dev} A| - b)/(2\mu + \eta). \tag{7.9}$$

Using this in (7.8) we deduce

$$P = \frac{(|\operatorname{dev} A| - b)_+}{2\mu + \eta} \frac{\operatorname{dev} A}{|\operatorname{dev} A|}. \tag{7.10}$$

The formula (7.10) holds also for $P = 0$. Taking (7.10) in (7.2) we calculate the minimal value (7.4). □

DEFINITION 7.1. For any $x \in \Omega$, $K(x) := \operatorname{dev} \mathbb{C}p_0(x) + \operatorname{dev} M(x)$, $M(x)$, $m(x)$ and η from Table 1, and $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ let

$$\varphi(x, A) := \frac{1}{2} A : \mathbb{C}A - \frac{1}{2}(|\operatorname{dev} \mathbb{C}A - K(x)| - m(x))_+^2/(2\mu + \eta).$$

PROPOSITION 7.2. For any $x \in \Omega$, $\varphi(x, \cdot)$ *is* $\mathscr{C}^1$ *and* $D\varphi(x, \cdot) : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to R^{d\times d}_{\mathrm{sym}}$ *is uniformly convex and Lipschitz, i.e. there exist positive hardening depending constants* α *and* L *such that, for all* $x \in \Omega$ *and* $A, B \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$,

$$\alpha|A - B|^2 + D\varphi(x, A) : (B - A) \leqslant \varphi(x, B) - \varphi(x, A), \tag{7.11}$$

$$|D\varphi(x, B) - D\varphi(x, A)| \leqslant L|B - A|. \tag{7.12}$$

PROOF. A discussion of the two cases $|\operatorname{dev} \mathbb{C}A - K(x)| \leq m(x)$ or not leads to

$$D\varphi(x, A) = \mathbb{C}\left(A - \frac{(|\operatorname{dev} \mathbb{C}A - K(x)| - m(x))_+}{(2\mu + \eta)} \frac{\operatorname{dev} \mathbb{C}A - K(x)}{|\operatorname{dev} \mathbb{C}A - K(x)|}\right) \tag{7.13}$$

and, at least for the two cases $|\operatorname{dev} \mathbb{C}A - K(x)| < m(x)$ resp. $> m(x)$, we have $D^2\varphi(x, A) = \mathbb{C}$ resp.

$$D^2\varphi(x, A) = \mathbb{C} - \frac{1 - m(x)/|B|}{2\mu + \eta} \mathbb{C} \operatorname{dev} \otimes \mathbb{C} \operatorname{dev} - \frac{m(x)}{(2\mu + \eta)|B|} \mathbb{C} \operatorname{sign} B \otimes \mathbb{C} \operatorname{sign} B\,, \tag{7.14}$$

where $B := \operatorname{dev} \mathbb{C}A - K(x)$, $\operatorname{sign} B = B/|B|$, and $P : (\mathbb{C} \operatorname{dev} \otimes \mathbb{C} \operatorname{dev})Q = 4\mu^2 \operatorname{dev} P : \operatorname{dev} Q$ for all $P, Q \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Since $\eta > 0$, $|B| \geq m(x)$ in the second case, and $\mathbb{C} \operatorname{dev} = 2\mu \operatorname{dev}$,

$$\frac{\eta}{\eta + 2\mu} \mathbb{C} \leq q\mathbb{C} - \frac{4\mu^2}{2\mu + \eta} \leq D^2\varphi(x, A) \leq \mathbb{C}\,, \tag{7.15}$$

where $\mathbb{A} \leq \mathbb{B}$ means $P : \mathbb{A}P \leq P : \mathbb{B}P$ for all $P \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$. Hence, the second derivative $D^2\varphi(x, A)$ is discontinuous but bounded. Then, (7.11) and (7.12) follow from integrations along lines connecting A and B; we refer to [20] for details. □

According to Definition 7.1 and Proposition 7.2 we may consider an equivalent variational problem.

DEFINITION 7.2 (Minimisation Problem (M_3)). Seek the minimiser u in $H^1_{\mathrm{D}}(\Omega)$ of

$$\int_\Omega \varphi(x, \varepsilon(u))\,\mathrm{d}x - \int_\Omega (fu + p_0 : \mathbb{C}\varepsilon(u))\,\mathrm{d}x - \int_{\Gamma_{\mathrm{N}}} gu\,\mathrm{d}s\,. \tag{7.16}$$

THEOREM 7.1. For $\eta > 0$, there exists a unique minimiser u of (7.16) and u, $\sigma := D\varphi(x, \varepsilon(u))$, and $p := \varepsilon(u) - \mathbb{C}^{-1}\sigma$ solve Problem (P_1), (P_2) and (M_1).

PROOF. Substituting Definition 7.1 for $A = \operatorname{dev} \mathbb{C}\varepsilon(U)$ into (7.16) and using (7.2) and (7.3) we eventually obtain (M_2). The unique solvability of (M_2) follows from standard arguments for uniformly monotone operators (see e.g. [20]). □

On the discrete level, we replace $m(x)$ and $K(x)$ by their elementwise integral means and result in the following discrete problem.

DEFINITION (Minimisation Problem $(M_3 \mid \mathscr{H})$). Seek the minimiser U in $\mathscr{H}$ of

$$\int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U))\,\mathrm{d}x - \int_\Omega (fU + \overline{p_0} : \mathbb{C}\varepsilon(U))\,\mathrm{d}x - \int_{\Gamma_{\mathrm{N}}} gU\,\mathrm{d}s \tag{7.17}$$

where, for $x \in T \in \mathscr{T}$ and $A \in \mathbb{R}^{d\times d}$,

$$\varphi_T(x, A) := \frac{1}{2} A : \mathbb{C}A - \frac{1}{2} (|\operatorname{dev} \mathbb{C}A - \overline{K}(x)| - \overline{m}(x))^2_+ / (2\mu + \eta)$$

and $\overline{K} := \int_T K(x)\,\mathrm{d}x/\operatorname{meas}(T)$ resp. $\overline{m} := \int_T m(x)\,\mathrm{d}x/\operatorname{meas}(T)$ denote integral means over T.

REMARK 7.1. The discrete problem $(P_2 \mid \mathscr{H} \times \mathscr{L})$ is equivalent to the minimisation problem $(M_3 \mid \mathscr{H})$.

REMARK 7.2. The functional $\varphi_{\mathscr{T}}$ satisfies (7.11) and (7.12) with the same constants. In particular, an a priori and a posteriori error analysis can also be based on these properties (but includes some implicit perturbation argument on $K - \overline{K}$, etc.).

DEFINITION 7.4 (Quasi–Newton–Raphson Scheme). Let $\mathbb{C}_n(x)$, $x \in \Omega$, be a globally elliptic and bounded fourth-order tensor, i.e. for two positive constants α_n and $\overline{\alpha_n}$ we have

$$\underline{\alpha_n}|A|^2 \leq A : \mathbb{C}_n(x)A \leq \overline{\alpha_n}|A|^2 \quad (A \in \mathbb{R}^{d\times d}_{\mathrm{sym}},\ x \in \Omega)\,. \tag{7.18}$$

Then, for $\rho_n > 0$ and $U_n \in \mathscr{H}$, define $U_{n+1} \in \mathscr{H}$ as the unique solution to

$$\int_\Omega \varepsilon(U_{n+1} - U_n) : \mathbb{C}_n \varepsilon(V)\,\mathrm{d}x = -\rho_n \Big\{ \int_\Omega D\varphi_{\mathscr{T}}(x, \varepsilon(U_n)) : \varepsilon(V)\,\mathrm{d}x - \int_\Omega (fV + \overline{p_0} : \mathbb{C}\varepsilon(V))\,\mathrm{d}x - \int_{\Gamma_N} gV\,\mathrm{d}s \Big\} \quad (V \in \mathscr{H})\,. \tag{7.19}$$

We have the following global, but only linear convergence result.

THEOREM 7.2. Let α and L as in Proposition 7.2. If $\rho_n = \underline{\alpha_n}/L$ then, any sequence (U_n) generated by the quasi-Newton–Raphson scheme satisfies

$$\frac{\alpha}{2(1+c)} \|\varepsilon(U - U_{n+1})\|^2_{L^2(\Omega)} + \delta_{n+1} \leq q\delta_n \tag{7.20}$$

for $q := c/(1+c) < 1$ with $1/2(1 + \overline{\alpha_n}/\underline{\alpha_n})^2 L^2/\alpha^2 \leq c$ and

$$\delta_n := \int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U_n))\,\mathrm{d}x - \int_\Omega (fU_n + p_0 : \mathbb{C}\varepsilon(U_n))\,\mathrm{d}x - \int_{\Gamma_N} gU_n\,\mathrm{d}s - \int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U))\,\mathrm{d}x + \int_\Omega (fU + \overline{p_0} : \mathbb{C}\varepsilon(U))\,\mathrm{d}x + \int_{\Gamma_N} gU\,\mathrm{d}s > 0\,. \tag{7.21}$$

PROOF. The proof follows from standard arguments given, e.g. in [20] for an abstract framework; we give a proof here for the convenience of the reader. According to (7.11) we have, for $e_n := \|\varepsilon(U - U_n)\|_{L^2(\Omega)}$,

$$\begin{aligned} \alpha e^2_{n+1} + \int_\Omega D\varphi_{\mathscr{T}}(x, \varepsilon(U_{n+1})) : \varepsilon(U - U_{n+1})\,\mathrm{d}x &\leq \int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U))\,\mathrm{d}x - \int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U_{n+1}))\,\mathrm{d}x \\ &= -\delta_{n+1} + \int_\Omega (f(U - U_{n+1}) + p_0 : \mathbb{C}\varepsilon(U - U_{n+1}))\,\mathrm{d}x + \int_{\Gamma_N} g(U - U_{n+1})\,\mathrm{d}s\,. \end{aligned} \tag{7.22}$$

Owing to (7.12), this shows

$$\begin{aligned} \alpha e^2_{n+1} + \delta_{n+1} \leq{}& \int_\Omega D\varphi_{\mathscr{T}}(x, \varepsilon(U_n)) : \varepsilon(U_{n+1} - U)\,\mathrm{d}x \\ &- \int_\Omega f(U_{n+1} - U) + p_0 : \mathbb{C}\varepsilon(U_{n+1} - U))\,\mathrm{d}x - \int_{\Gamma_N} g(U_{n+1} - U)\,\mathrm{d}s \\ &+ L\|\varepsilon(U_{n+1} - U)\|_{L^2(\Omega)}\|\varepsilon(U_{n+1} - U_n)\|_{L^2(\Omega)}\,. \end{aligned} \tag{7.23}$$

Taking (7.19) for $V = U - U_{n+1}$, we obtain in (7.23) that

$$\begin{aligned} \alpha e^2_{n+1} + \delta_{n+1} &\leq \frac{1}{\rho_n} \int_\Omega \varepsilon(U_{n+1} - U_n) : \mathbb{C}_n \varepsilon(U - U_{n+1})\,\mathrm{d}x + L\|\varepsilon(U_{n+1} - U)\|_{L^2(\Omega)}\|\varepsilon(U_{n+1} - U_n)\|_{L^2(\Omega)} \\ &\leq (L + \overline{\alpha_n}/\rho_n)\|\varepsilon(U_{n+1} - U_n)\|_{L^2(\Omega)}\|\varepsilon(U_{n+1} - U)\|_{L^2(\Omega)}\,. \end{aligned} \tag{7.24}$$

(In the last step we used Cauchy's inequality with respect to $\mathbb{C}_n$ and (7.18).) In the same way, we apply (7.11), (7.12) and (7.19), to infer

$$\begin{aligned} \alpha\|\varepsilon(U_{n+1} - U_n)\|^2_{L^2(\Omega)} &\leq \int_\Omega D\varphi_{\mathscr{T}}(x, \varepsilon(U_{n+1})) : \varepsilon(U_{n+1} - U_n)\,\mathrm{d}x \\ &\quad + \int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U_n))\,\mathrm{d}x - \int_\Omega \varphi_{\mathscr{T}}(x, \varepsilon(U_{n+1}))\,\mathrm{d}x \\ &= \delta_n - \delta_{n+1} + \int_\Omega (f(U_n - U_{n+1}) + p_0 : \mathbb{C}\varepsilon(U_n - U_{n+1}))\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&+\int_{\Gamma_N} g(U_n - U_{n+1})\,ds + \int_\Omega D\varphi_{\mathcal{F}}(x, \varepsilon(U_{n-1}) : \varepsilon(U_{n+1} - U_n)\,dx \\
&= \delta_n - \delta_{n+1} - \frac{1}{\rho_n}\int_\Omega \varepsilon(U_{n+1} - U_n) : \mathbb{C}\varepsilon(U_{n+1} - U_n)\,dx \\
&\quad + \int_\Omega \{D\varphi_{\mathcal{F}}(x, \varepsilon(U_{n+1}) - D\varphi_{\mathcal{F}}(x, \varepsilon(U_n)\} : \varepsilon(U_{n+1} - U_n)\,dx \\
&\leq \delta_n - \delta_{n+1} - \underline{\alpha_n}/\rho_n \|\varepsilon(U_{n+1} - U_n)\|^2_{L^2(\Omega)} + L\|\varepsilon(U_{n+1} - U_n)\|^2_{L^2(\Omega)} \\
&= \delta_n - \delta_{n+1}
\end{aligned}
\tag{7.25}
$$

because of $\rho_n := \underline{\alpha_n}/L$. Using this in (7.24) with Young's inequality, $ab \leq a^2/2 + b^2/2$, we deduce

$$
\alpha e^2_{n+1} + \delta_{n+1} \leq \frac{\alpha}{2} e^2_{n+1} + \frac{L^2}{2\alpha^2}(1 + \overline{\alpha_n}/\underline{\alpha_n})^2(\delta_n - \delta_{n+1})\,. \tag{7.26}
$$

This shows (7.20) and concludes the proof. □

REMARK 7.3. The (non-damped) Newton–Raphson method is defined for $\rho_n = 1$ and $\mathbb{C}_n := D^2\varphi_{\mathcal{F}}(x, \varepsilon(U_n))$. Then, since $\alpha/L < 1$, Theorem 7.2 does not guarantee global convergence. Moreover, q may be very close to 1 if hardening is small.

We conclude this section with a discussion of the local convergence properties of the Newton–Raphson scheme to indicate superlinear convergence observed in practice. Suppose $U_n \in \mathcal{H}$ is an approximation to $U \in \mathcal{H}$ and $U_{n+1} \in \mathcal{H}$ is generated by

$$
\begin{aligned}
&\int_\Omega D^2\varphi_{\mathcal{F}}(x, \varepsilon(U_n); \varepsilon(U_{n+1} - U_n), \varepsilon(V))\,dx \\
&\quad = -\int_\Omega D\varphi_{\mathcal{F}}(x, \varepsilon(U_n); \varepsilon(V))\,dx + \int_\Omega (fV + p_0 : \mathbb{C}\varepsilon(V))\,dx + \int_{\Gamma_N} gV\,ds \quad (V \in \mathcal{H})\,.
\end{aligned}
\tag{7.27}
$$

With the (discrete) stress deviator resp. its approximation

$$
\sigma := \operatorname{dev}\mathbb{C}[\varepsilon(U) - \overline{K}] \quad \text{resp.} \quad \sigma_n := \operatorname{dev}\mathbb{C}[\varepsilon(U_n) - \overline{K}] \tag{7.28}
$$

we define the discrete plastic (and elastic) zone resp. its approximation by

$$
\underset{(e)}{\Omega_p} := \{x \in \Omega \,|\, |\sigma(x)| \underset{(>)}{\leq} \overline{m}(x)\} \text{ resp. } \underset{(e)}{\Omega_{n,p}} := \{x \in \Omega \,|\, |\sigma_n(x)| \underset{(>)}{\leq} \overline{m}(x)\}\,. \tag{7.29}
$$

Thus, the prediction of U_n about the plastic or elastic zone is correct on

$$
(\Omega_p \cap \Omega_{n,p}) \cup (\Omega_e \cap \Omega_{n,e}) \tag{7.30}
$$

and incorrect on the remaining part of $\Omega_{n,i} := (\Omega_p \backslash \Omega_{n,p}) \cup (\Omega_e \backslash \Omega_{n,e})$ of the domain. We will work under the hypothesis that, first, $\|\varepsilon(U - U_n)\|_{L^\infty(\Omega)}$ is small and, second, meas $(\Omega_{n,i})$ is very small. Our analysis is based on the fact that

$$
d_n(x) := D^2\varphi_{\mathcal{F}}(x, \varepsilon(U_n)(x); \varepsilon(U - U_n)(x)) + D\varphi_{\mathcal{F}}(x, \varepsilon(U_n)(x)) - D\varphi_{\mathcal{F}}(x, \varepsilon(U)(x)) \tag{7.31}
$$

and (7.27) lead, for $V := U_{n+1} - U$, to

$$
\int_\Omega D^2\varphi_{\mathcal{F}}(x, \varepsilon(U_n); \varepsilon(U_{n+1} - U), \varepsilon(U_{n+1} - U))\,dx = -\int_\Omega d_n : \varepsilon(U_{n+1} - U)\,dx\,. \tag{7.32}
$$

This, Cauchy's inequality and the uniform positive definiteness of $D^2\varphi_{\mathcal{F}}$ from (7.15) yield, with some η-depending constant $c_1(\eta) > 0$, that

$$
\|\varepsilon(U - U_{n+1})\|_{L^2(\Omega)} \leq c_1(\eta)\|d_n\|_{L^2(\Omega)} \tag{7.33}
$$

and it remains to analyse (7.31). Recalling the proof of Proposition 7.2 we find

$$|d_n(x)| \leq c_2(\eta)|\varepsilon(U - U_n)(x)| \quad \text{for almost all } x \in \Omega\,, \tag{7.34}$$

since $D^2\varphi_{\mathcal{F}}$ is bounded, $D\varphi_{\mathcal{F}}$ is Lipschitz. Furthermore, $D^2\varphi_{\mathcal{F}}(\cdot, \varepsilon(U_n))$ is constant $\mathbb{C}$ in $\Omega_{n,e}$, and so

$$d_n(x) = 0 \quad \text{for almost all } x \in \Omega_e \cap \Omega_{n,e}\,. \tag{7.35}$$

Next, consider $x \in \Omega_p \cap \Omega_{n,p}$ and $A = \varepsilon(U_n)(x)$ and $B = \varepsilon(U)(x)$. We have $A, B \in \{C \in \mathbb{R}^{d\times d}_{\mathrm{sym}} \,|\, \mathrm{dev}\,\mathbb{C}[C - \overline{K}(x)] > \overline{m}(x)\} =: \mathcal{P}$ and, although $\mathrm{conv}\{A, B\}$ may leave $\mathcal{P}$, we can always find a smooth curve $\gamma : [0, l] \to \mathcal{P}$ with $\gamma(0) = A$ and $\gamma(l) = B$ of length

$$l \leq c_3|A - B|\,. \tag{7.36}$$

Since $D^2\varphi_{\mathcal{F}}$ is smooth in $\mathcal{P}$ and $D^3\varphi_{\mathcal{F}}$ is bounded (cf. the derivative of (7.14)), we obtain (with γ parameterised by the arc-length, and so $|\dot\gamma(s)| = 1$)

$$\begin{aligned}
d_n(x) &= D^2\varphi_{\mathcal{F}}(A; B - A) + D\varphi_{\mathcal{F}}(A) - D\varphi_{\mathcal{F}}(B)\\
&= D^2\varphi_{\mathcal{F}}(A; B - A) - \int_0^l \frac{\partial}{\partial s} D\varphi_{\mathcal{F}}(\gamma(s))\,\mathrm{d}s\\
&= \int_0^l \{D^2\varphi_{\mathcal{F}}(A; \dot\gamma(s)) - D^2\varphi(\gamma(s); \dot\gamma(s))\}\,\mathrm{d}s\\
&\leq \int_0^l |D^2\varphi_{\mathcal{F}}(A) - D^2\varphi_{\mathcal{F}}(\gamma(s))|\,\mathrm{d}s\\
&= \int_0^l \left| \int_0^s \frac{\partial}{\partial t} D^2\varphi_{\mathcal{F}}(\gamma(t))\,\mathrm{d}t \right| \mathrm{d}s \leq \int_0^l \int_0^s |D^3\varphi_{\mathcal{F}}(\gamma(t))|\,\mathrm{d}t\,\mathrm{d}s\,.
\end{aligned} \tag{7.37}$$

According to the boundedness of $D^3\varphi$ and (7.36), (7.37) verifies

$$|d_n(x)| \leq c_4(\eta)|\varepsilon(U - U_n)(x)|^2 \quad \text{for almost all } x \in \Omega_p \cap \Omega_{n,p}\,. \tag{7.38}$$

To summarise, the above arguments show

$$\begin{aligned}
\|\varepsilon(U - U_{n+1})\|_{L^2(\Omega)} &\leq 2c_2(\eta)\|\varepsilon(U - U_n)\|_{L^2(\Omega_{n,i})}\\
&\quad + 2c_4(\eta)\|\varepsilon(U - U_n)\|_{L^\infty(\Omega_p\cap\Omega_{n,p})}\|\varepsilon(U - U_n)\|_{L^2(\Omega_p\cap\Omega_{n,p})}\,.
\end{aligned} \tag{7.39}$$

Thus, if $\|\varepsilon(U - U_n)\|_{L^\infty(\Omega)} \leq \delta \ll 1$ and $\mathrm{meas}(\Omega_{n,i}) \leq \delta^{2\alpha}$, with $0 < \alpha \leq 1$, then,

$$\|\varepsilon(U - U_{n+1})\|_{L^2(\Omega)} = 0(\delta^{1+\alpha}) \quad (\delta \to 0)\,. \tag{7.40}$$

REMARK 7.4. Under the present assumptions on $\varepsilon(U - U_n)$ we can, in general, only expect $\|\varepsilon(U - U_n)\|_{L^2(\Omega)} = 0(\delta)$, and so (7.40) indicates superlinear convergence of Newton–Raphson's scheme.

REMARK 7.5. Although (7.40) may suggest that the local convergence of Newton–Raphson's method is superlinear, it seems to be a non-obvious task to base a rigorous proof on (7.38). The difficulty arises from the fact that, within a proof by mathematical induction, we have to verify that $\|\varepsilon(U - U_{n+1})\|_{L^\infty(\Omega)} = \mathrm{O}(\delta^{1+\alpha})$ and $\mathrm{meas}(\Omega_{n+1,i}) = \mathrm{O}(\delta^{2\alpha(1+\alpha)})$.

REMARK 7.6. In our practical experience, the Newton–Raphson scheme often showed quadratic convergence.

REMARK 7.7. It is conjectured that damping is necessary in the first steps of the iteration and may be omitted in subsequent iterations.

8. Numerical experiments

Three examples provide numerical evidence of the linear convergence of the lowest order scheme and of the superiority of the adaptive algorithm.

8.1. Plastic ring with known solution

The first example involves kinematic hardening for the geometry shown in Fig. 1 which represents the two-dimensional section of a long tube with inner radius of 1 and an outer radius of 2, where we have no volume force ($f=0$) but radially applied surface forces $g_1(r,\phi,t)=te_r$ and $g_2(r,\phi,t)=-t/4e_r$, $e_r=(\cos\phi,\sin\phi)$.

The lack of Dirichlet boundary conditions is compensated by requiring that Ω keeps centred at the origin and rotation is prohibited. The rotational invariant system and applied loads allow the solution

$$u(r,\phi,t)=u_r(r,t)\cdot e_r,$$
$$\sigma(r,\phi,t)=\sigma_r(r,t)e_r\otimes e_r+\sigma_\phi(r,t)e_\phi\otimes e_\phi,$$
$$p(r,\phi,t)=P_r(r,t)\cdot(e_r\otimes e_r-e_\phi\otimes e_\phi),$$

(see [21] for details) with $e_r=(\cos\phi,\sin\phi)$, $e_\phi=(-\sin\phi,\cos\phi)$ and

$$u_r(r,t)=\begin{cases}\dfrac{t}{2\mu r}-\dfrac{1}{3}\kappa I(R(t))\left(r+\dfrac{4a}{\mu r}\right) & \text{for } r\geq R(t),\\ \dfrac{t}{2\mu r}-\dfrac{1}{3}\kappa I(R(t))\left(4r+\dfrac{4a}{\mu r}\right)+\kappa I(r)\cdot r & \text{for } r<R(t),\end{cases}$$

$$\sigma_r(r,t)=\begin{cases}-\dfrac{t}{r^2}-\dfrac{8}{3}a\kappa I(R(t))\left(\dfrac{1}{4}-\dfrac{1}{r^2}\right) & \text{for } r\geq R(t),\\ -\dfrac{t}{r^2}-\dfrac{8}{3}a\kappa I(R(t))\left(1-\dfrac{1}{r^2}\right)+2a\kappa I(r) & \text{for } r<R(t),\end{cases}$$

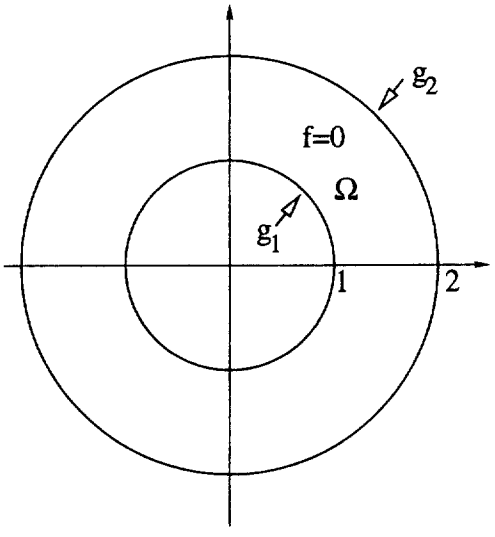


Fig. 1. Geometry in Example 8.1.

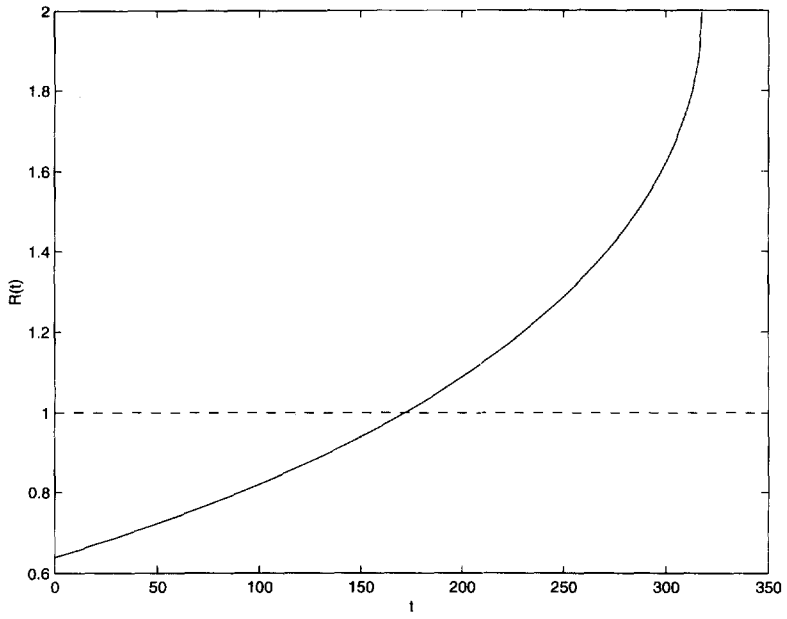


Fig. 2. Evolution in time of the plastic boundary $R(t)$ in Example 8.1.

$$\sigma_\phi(r,t) = \partial(r\cdot\sigma_r)/\partial r\,,$$

$$P_r(r,t) = \begin{cases} 0 & \text{for } r \geqslant R(t)\,, \\ \dfrac{\sigma_y}{\sqrt{2}(a\kappa + H_2)}\,(1 - R^2/r^2) & \text{for } r < R(t)\,, \end{cases}$$

$$I(r) = \frac{\sigma_y}{\sqrt{2}(a\kappa + H_2)}\,(\ln r + 1/2(R^2/r^2 - R^2))\,.$$

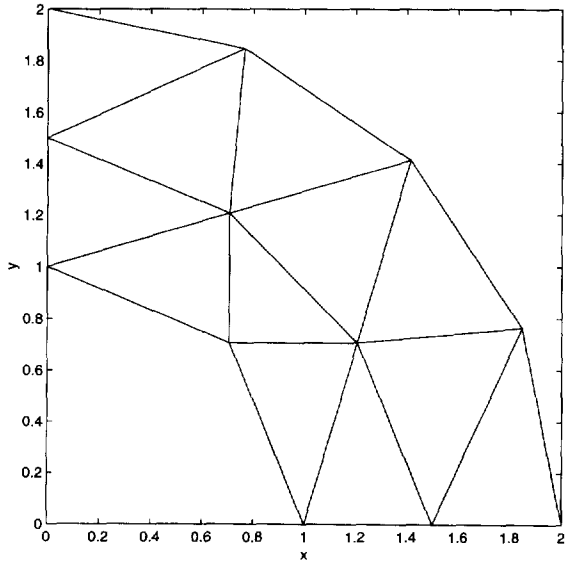


Fig. 3. Initial mesh in Example 8.1.

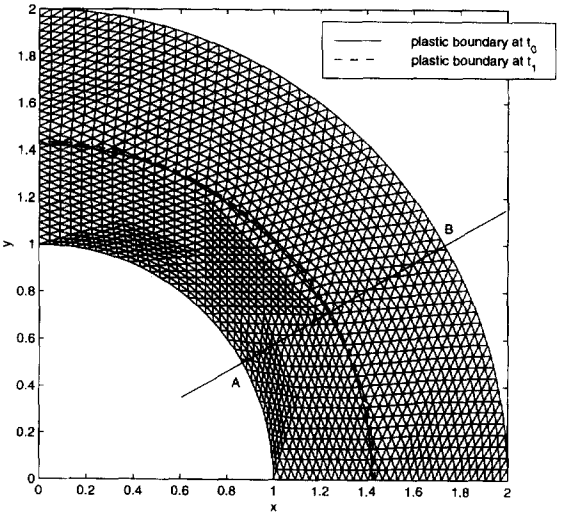
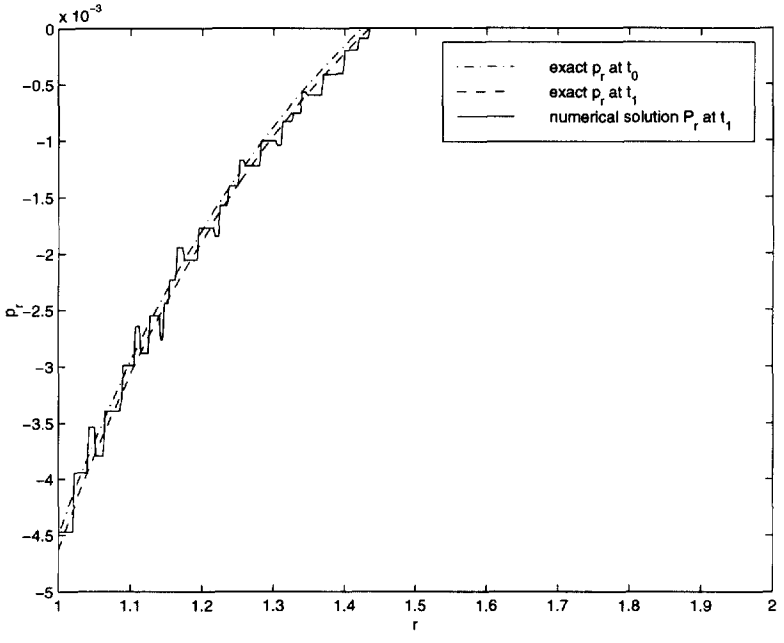


Fig. 4. Uniform mesh after 4 red-refinements in Example 8.1.

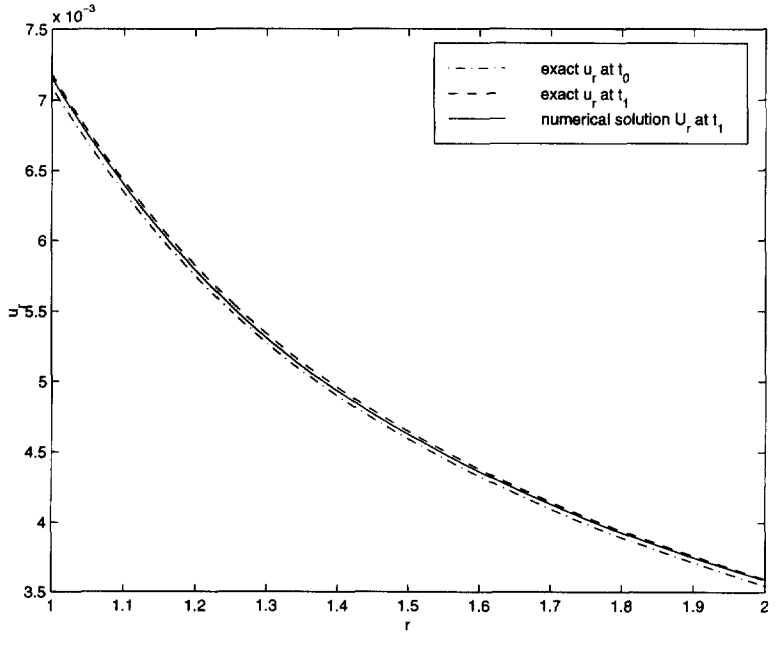
Here, $a = \mu + \lambda$, $\kappa = 2\mu/(2\mu + \lambda)$. The radius of the circular plastic boundary $R(t)$ is determined as the positive root of

$$f(R) = -2\alpha \ln R + (\alpha - 1)R^2 - \alpha + \frac{\sqrt{2}}{\sigma_y} t,$$

where $\alpha = 4a\kappa/(3(a\kappa + H_2))$. Fig. 2 displays $R(t)$ versus t and illustrates for $t \leq \sigma_y/\sqrt{2}$ that the body reacts purely elastic (as the inner radius of the domain is 1).



(a)

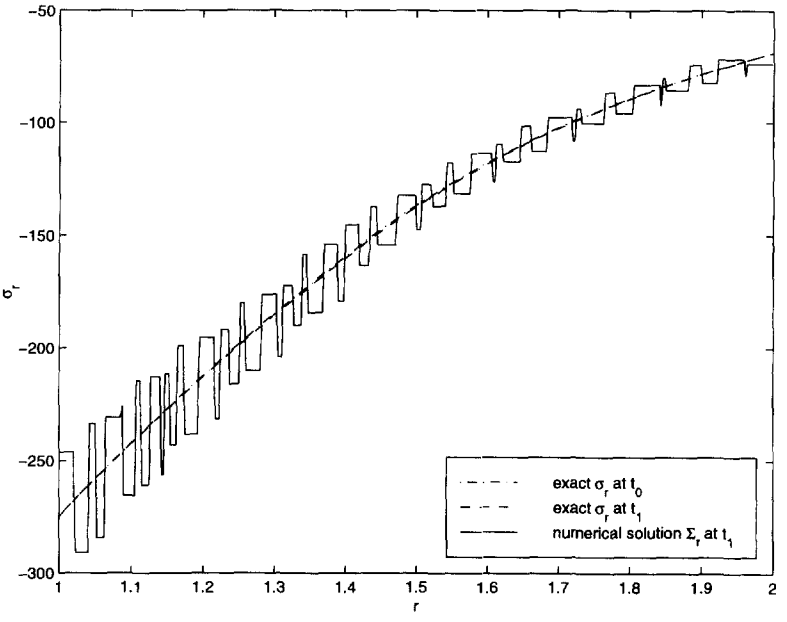


(b)

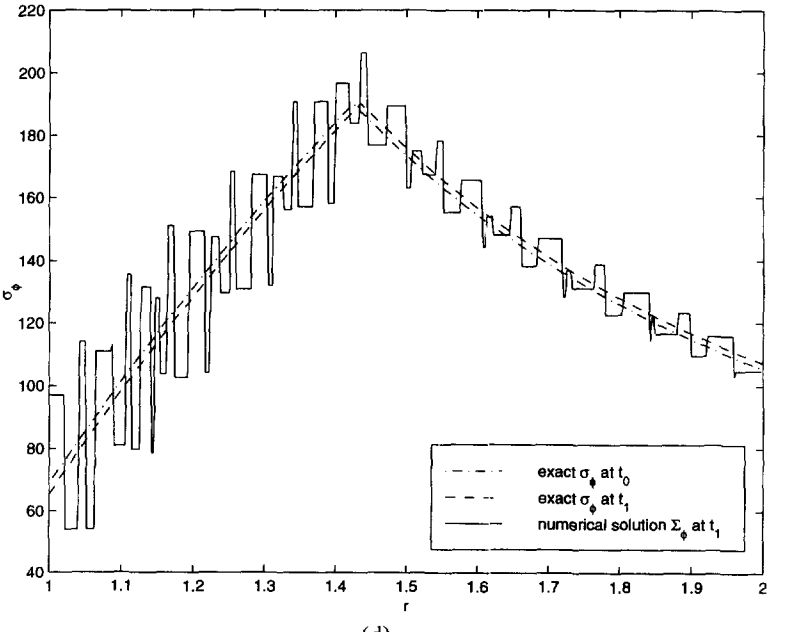
Fig. 5. Exact and discrete solution for the mesh shown in Fig. 4 along A–B in Example 8.1.

For material parameters with Lamé constants $\mu = E/(2(1+\nu))$ and $\lambda = \nu E/((1+\nu)(1-2\nu))$, Young's modulus $E = 70\,000$ and Poisson's ratio $\nu = 0.33$, yield stress $\sigma_y = 243$ and hardening modulus $\eta = 1$ for kinematic hardening, the inner part of the body becomes plastic at $t = 171.8269$. As a typical time-step we realized the time-increment from $t_0 = 274.9231$ to $t_1 = 276.6414$. According to the symmetry, only a quarter of the domain is discretized as shown in Fig. 3 with symmetric boundary conditions.

The mesh after 4 uniform refinement steps together with the boundary between the plastic and the elastic region at t_0 (solid line) and at t_1 is shown in Fig. 4. The new nodes on the boundary are moved to the curved boundary within each refinement step and so Fig. 4 indeed corresponds to red-refinements.



(c)



(d)

Fig. 5. (Contd.)

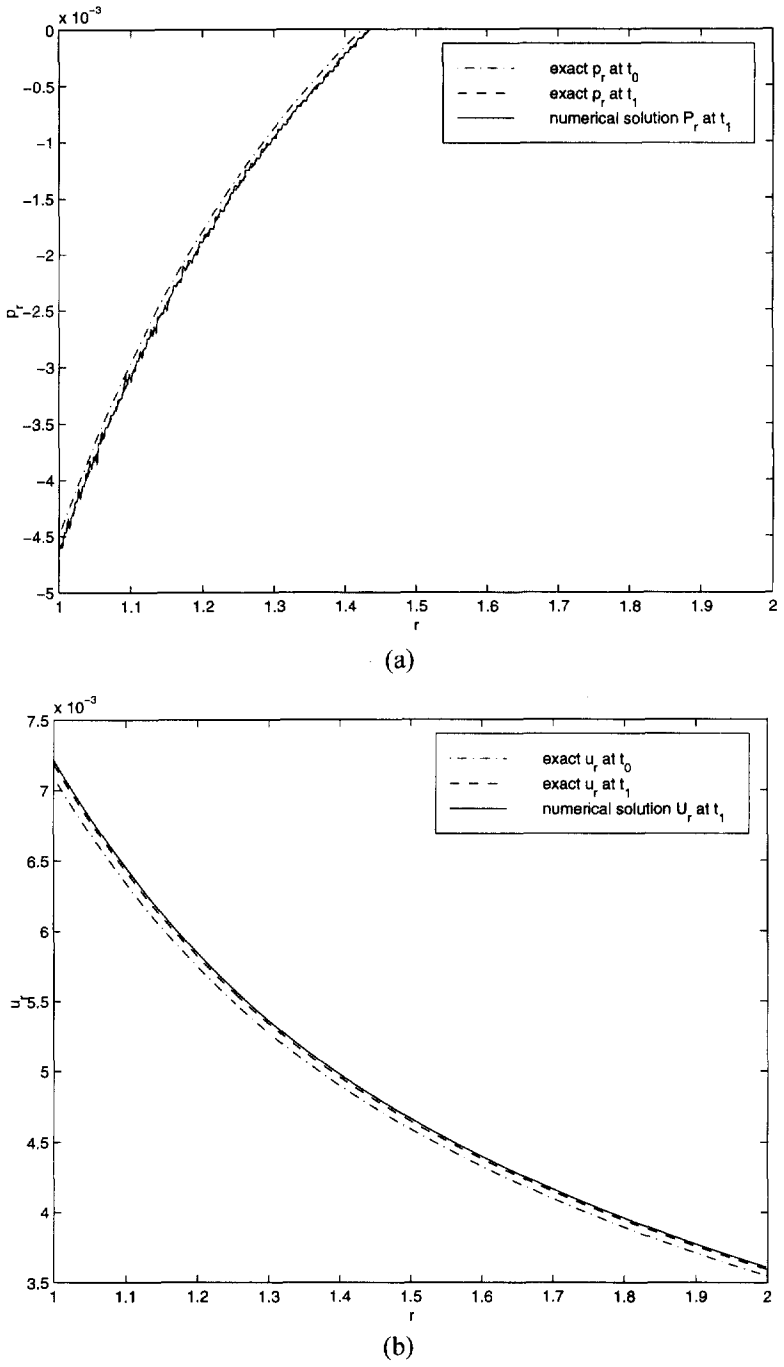
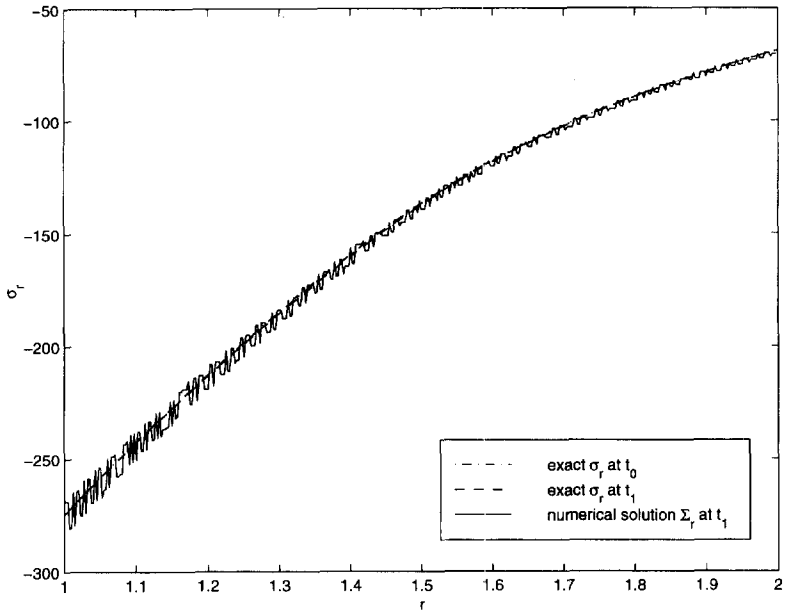


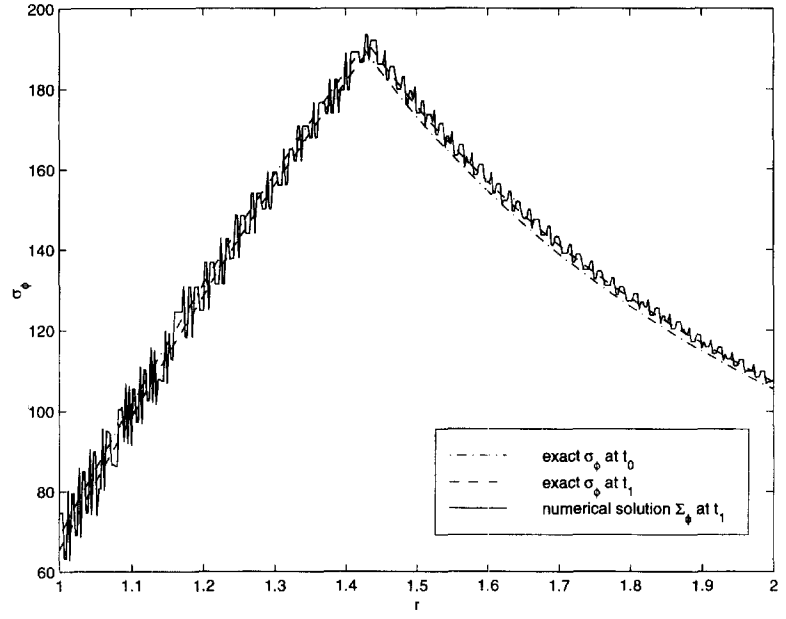
Fig. 6. Exact and discrete solution for the mesh shown in Fig. 7 along A–B in Example 8.1.

Along the line A–B the exact and the numerical solutions are compared in Figs. 5 and 6.

The figures show the radial components of the plastic strain p_r, the two components of the stress σ_r and σ_ϕ, and the displacement u_r, each for different mesh sizes. The error of the numerical solution seems to be uniformly distributed in radial direction in Figs. 5 and 6. Hence, no particular refinement of certain regions appears to be necessary and so the automatic mesh-refinement leads to a mostly uniformly refined mesh, shown in Fig. 7.



(c)



(d)

Fig. 6. (Contd.)

Finally, the inverse of the relative energy error in σ over the square root of the degrees of freedom n (as a measure for the mesh size) is displayed in Fig. 8 in a double logarithmic plot. The rate of convergence for the uniform and the adaptive mesh-refinement can be deduced from Fig. 8 and so provides experimental evidence that the asymptotic is indeed $O(h)$. Also, the convergence rate for the adaptive mesh–refinement with Algorithm 6.1 (dashed line) shows optimal convergence in agreement with the discussion above.

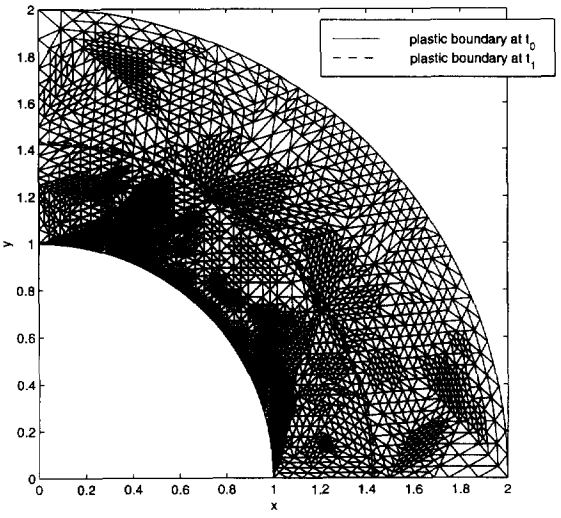


Fig. 7. Adaptively refined mesh in Example 8.1

8.2. Perforated tension strip

The second experiment is part of a benchmark in computational plasticity [22]. A two-dimensional squared plate with a hole is under a time-dependent tension $g(t) = 600 \cdot t$ with a surface load $g(t) = t$ as shown in Fig. 9. Only a quarter of the domain, depicted in Fig. 10, is discretized.

The problem models perfectly plastic material with Young's modulus $E = 206\,900$, Poisson's ratio $\nu = 0.29$, yield stress $\sigma_y = 450$, and with vanishing initial data for the plastic strain p_0 and hardening parameter α_0. The finite element method is applied to a uniformly refined mesh with 31 744 elements and $n = 32\,226$ degrees of freedom and a time integration from 0 to $T = 5$ with the implicit Euler method in 100 uniform time steps of length $k = 1/20$. In each time step, a Newton–Raphson method is employed which terminates if the Euclidean

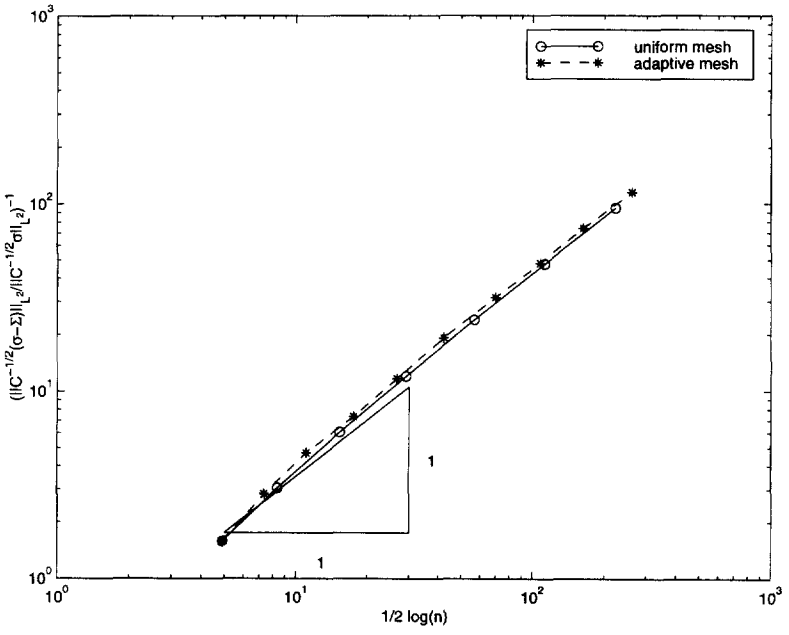


Fig. 8. Convergence rate in Example 8.1.

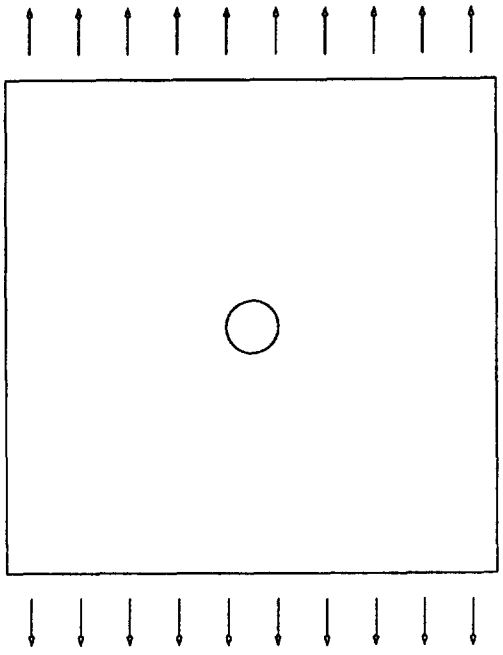


Fig. 9. Perforated tension strip in Example 8.2.

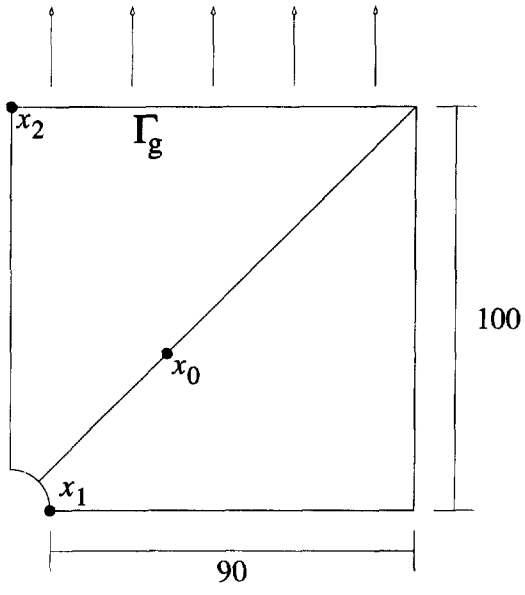


Fig. 10. Geometry of the problem in Example 8.2.

norm of the residual is less than 10^{-11}. In Fig. 11, we display the number of iterations needed in each time-step. Since the material remains entirely elastic for $0 \leq t < 1.75$, one iteration step is sufficient. The structure seems to fail for $t > 4.8$ where the plastic zone reaches the outer boundary and the Newton–Raphson scheme diverges. In Fig. 12 we display the two stress fields for time $t = 0.3$ resp. $t = 4.4$ in the deformed configuration (displacements are 10 times magnified), where the shading corresponds to $|\mathrm{dev}\,\Sigma|$. As a measure for the displacement on Γ_g, the curve in Fig. 13 shows the line integral of u_2 versus t. Again, failure of the structure is seen in the last four time-steps when the plastic region reaches Γ_g.

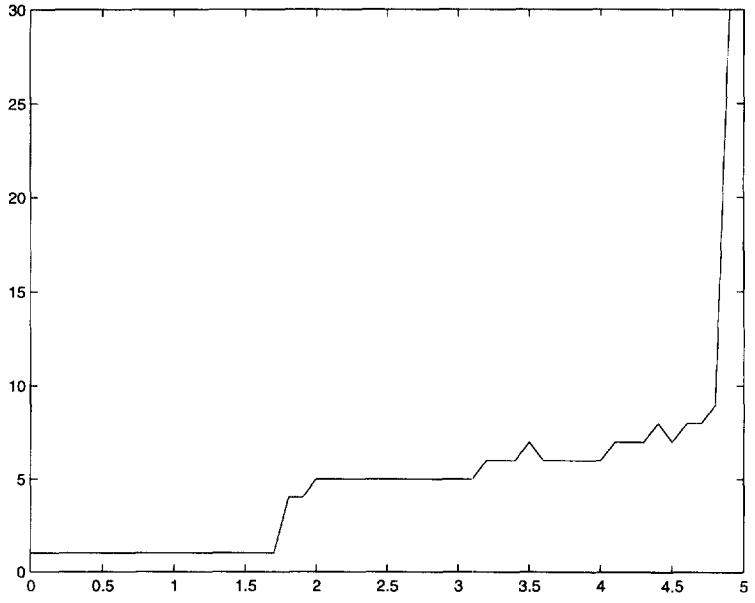


Fig. 11. Number of Newton iterates versus t, $0 \leq t \leq 5$ in Example 8.2.

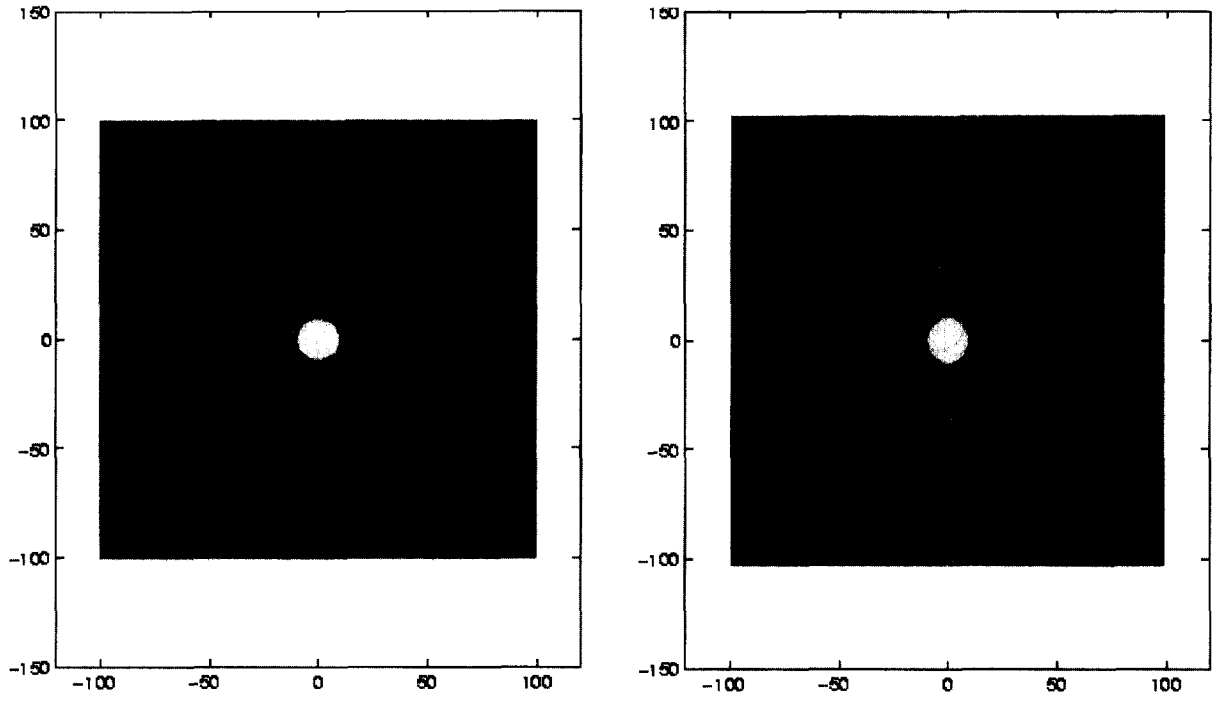


Fig. 12. Modulus of deviatoric part of stress approximation for $t = 0.3$ (left) and $t = 4.4$ (right) in Example 8.2.

Another quantity of interest is the first component of the stress tensor $\sigma_{11}(x_0)$, plotted in Fig. 14 (cf. Fig. 10 for the location of $x_0 = (38.044, 38.044)$).

8.3. Adaptive treatment within one time-step

In the final example, we continue the preceding benchmark and focus on one time-step, i.e. on Hencky

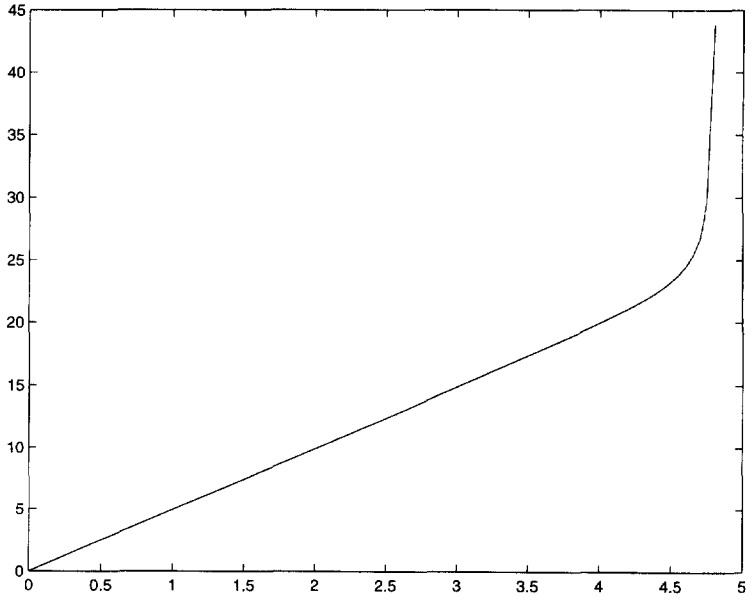


Fig. 13. $\int_{\Gamma_p} u_2 \, ds$ versus t in Example 8.2.

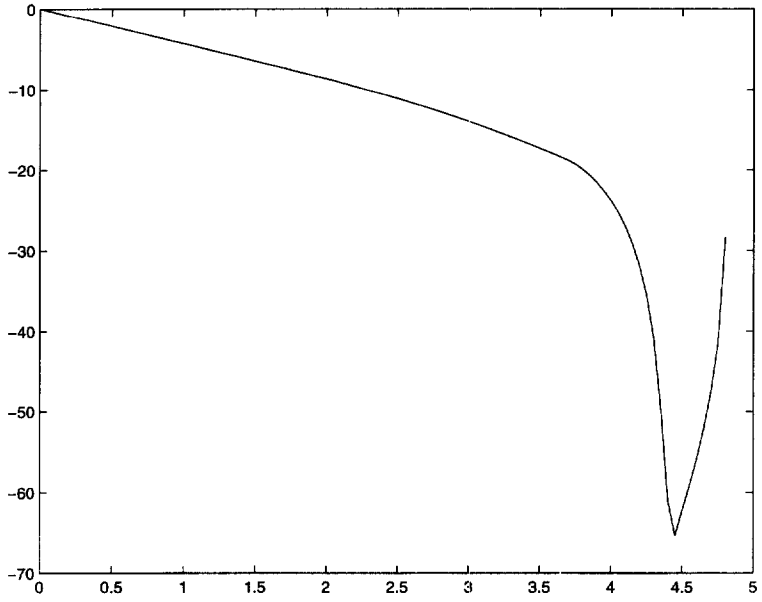


Fig. 14. $\sigma_{11}(x_0)$ versus t in Example 8.2.

plasticity, with a constant surface load $g = 450$ on Γ_g (see Fig. 10 for g and the location of Γ_g) and homogeneous initial data $p_0 = 0$ resp. $\alpha_0 = 0$.

Starting with the initial mesh T_0 plotted in Fig. 15, 30 adaptive steps of Algorithm 6.1 lead to a final configuration with $n = 8004$ degrees of freedom. The triangulations T_5, $\mathcal{T}_{10}$, $\mathcal{T}_{15}$, $\mathcal{T}_{20}$, $\mathcal{T}_{25}$, and $\mathcal{T}_{30}$ are displayed in Figs. 16–18.

Although the mesh $\mathcal{T}_0$ is a priori refined near the hole (to approximate properly the curved boundary there),

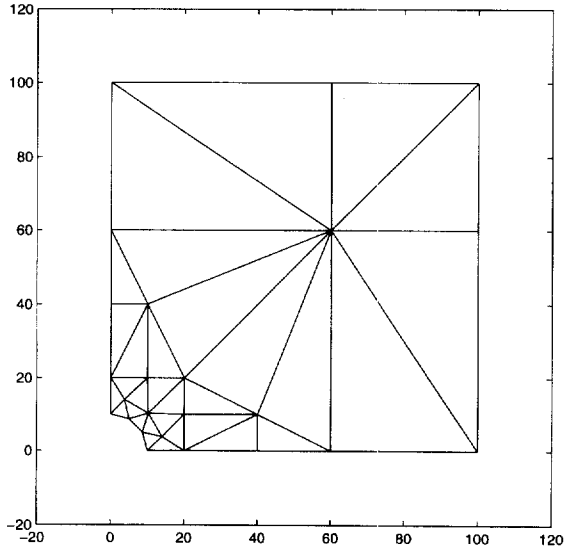


Fig. 15. Initial mesh $\mathcal{T}_0$ with $n = 48$ in Example 8.3.

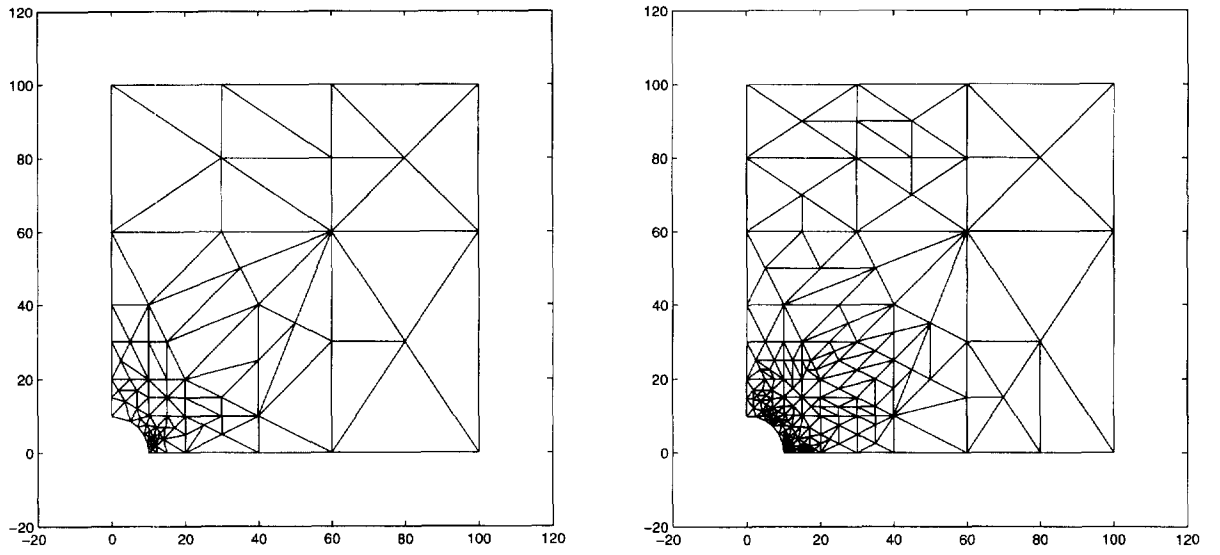


Fig. 16. $\mathcal{T}_5$ with $n = 150$ (left), $\mathcal{T}_{10}$ with $n = 340$ (right) in Example 8.3.

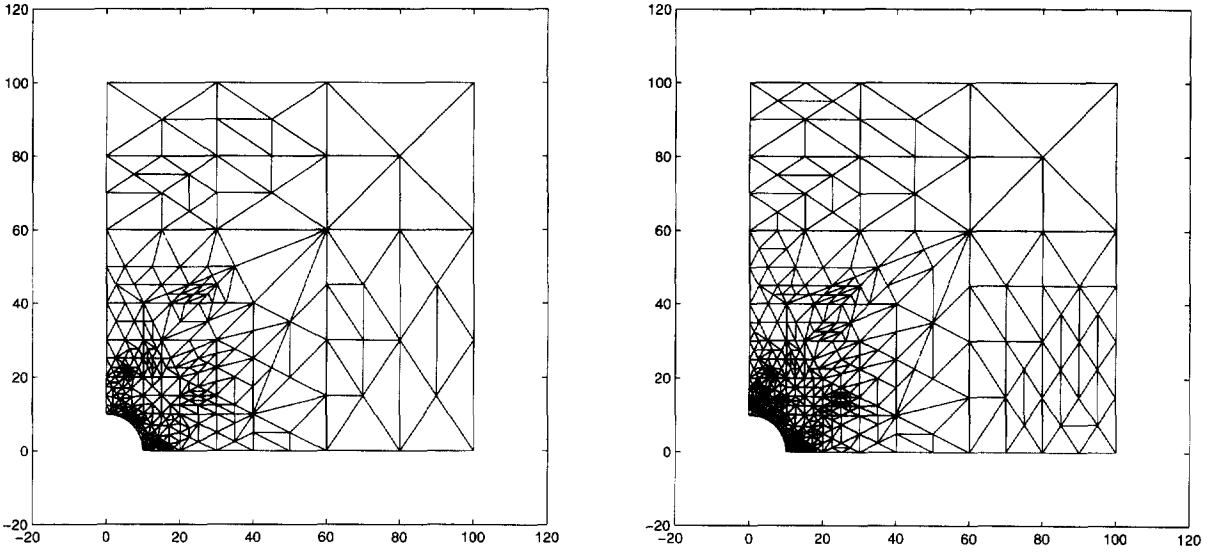


Fig. 17. $\mathcal{T}_{15}$ with $n = 544$ (left), $\mathcal{T}_{20}$ with $n = 854$ (right) in Example 8.3.

the refinement starts in the plastic region. The elastic part of the body is refined for very fine meshes $\mathcal{T}_k$. This behaviour is reflected by Fig. 19, where the stress component σ_{22} is evaluated at the point x_1 (see Fig. 10 for the location of x_1). The values oscillate up to a mesh with less than 4000 degrees of freedom and then reach a plateau. The values $\sigma_{22}(x_1)$ in the uniform refinement grow monotonously, but do not reach the approximate value $\sigma_{22}(x_1) = -8.6253$ (of the last adaptive step). Similar observations are made for the first component of the displacement at x_1 plotted in Fig. 20. Such oscillations of the (displacement and strain) values are not observed in the elastic regions of the domain in Figs. 21 and 22.

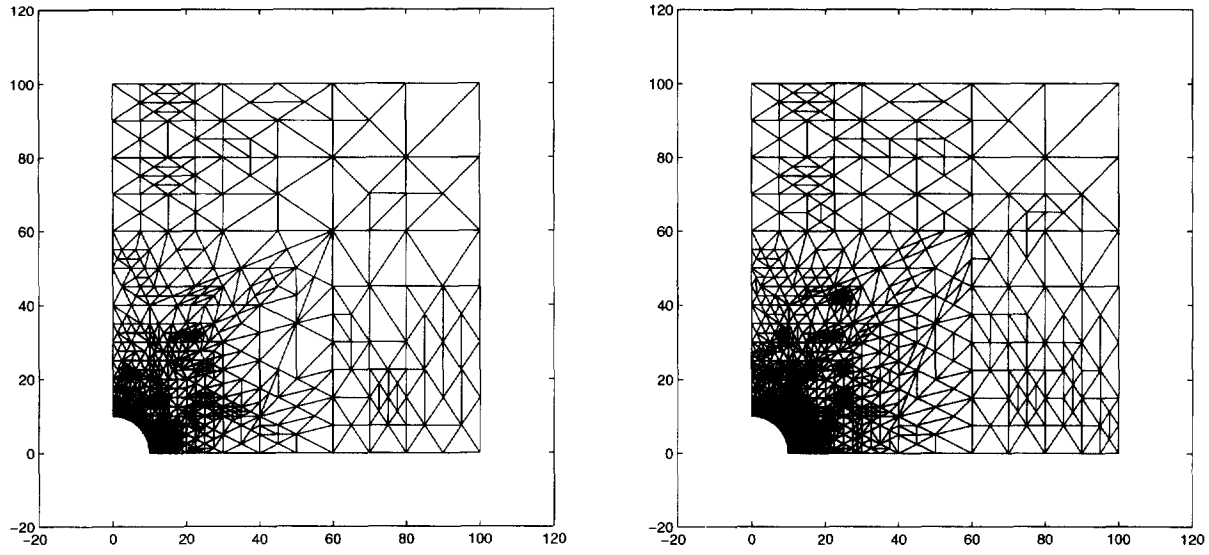


Fig. 18. $\mathcal{T}_{25}$ with $n = 1264$ (left), $\mathcal{T}_{30}$ with $n = 2030$ (right) in Example 8.3.

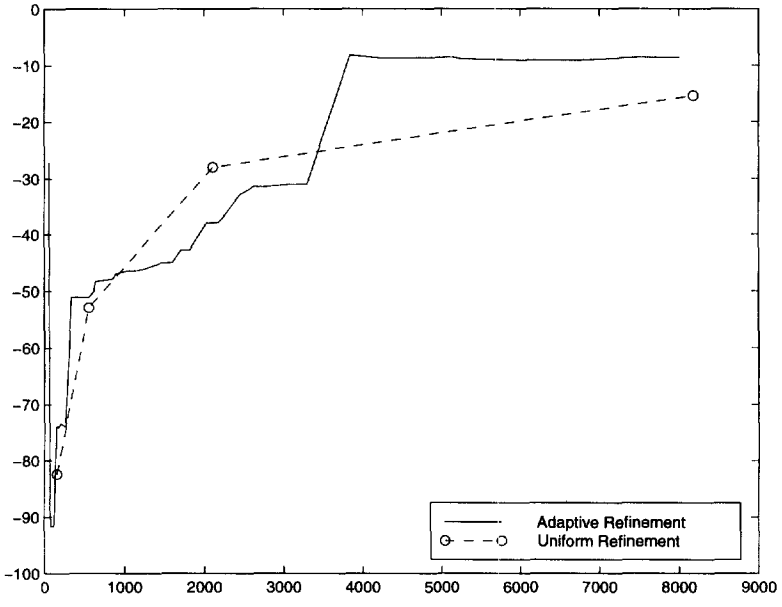


Fig. 19. $\sigma_{22}(x_1)$ versus degrees of freedom in Example 8.3.

To investigate the effectivity of the applied error indicator, we depicted the experimental convergence of the a posteriori error bound for approximations obtained with an adaptive mesh-refinement with Algorithm 6.1. In Fig. 23, we displayed the upper bound in (6.2),

$$\eta(\mathcal{T}_k) = \left(\sum_{T \in \mathcal{T}_k} \eta_T^2 \right)^{1/2}$$

with η_T from (6.1), versus the square root of the number of degrees of freedom $\sqrt{n}$ in $\mathcal{T}_k$ for ten sequences of

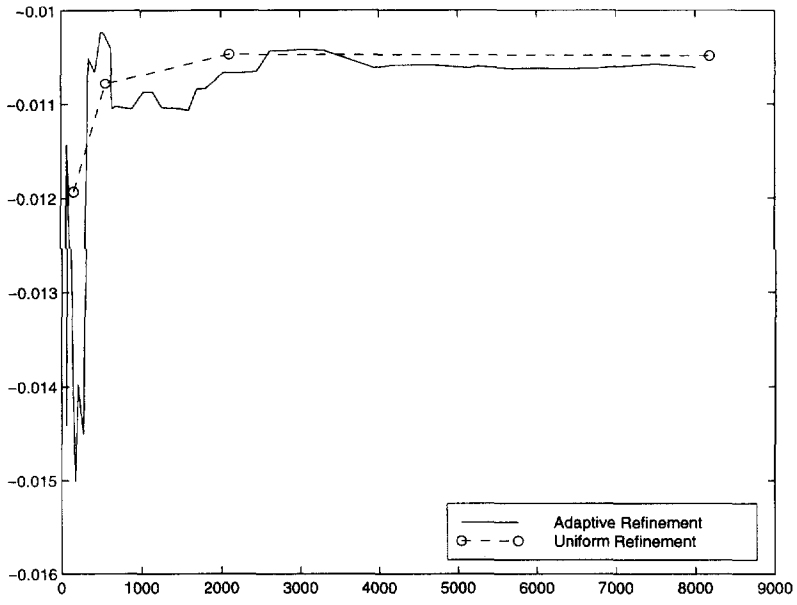


Fig. 20. $u_1(x_1)$ versus degrees of freedom in Example 8.3.

meshes generated adaptively ($k = 0, \ldots, 50$) corresponding to the depicted hardening parameters. The kinematic hardening leads to a relative raise of coercivity $\eta/(\eta + 2\mu)$. Although our a posteriori analysis is valid only for $\eta > 0$, in all six sequences we observe the optimal asymptotic slope 1 which corresponds to linear convergence of the a posteriori error bound.

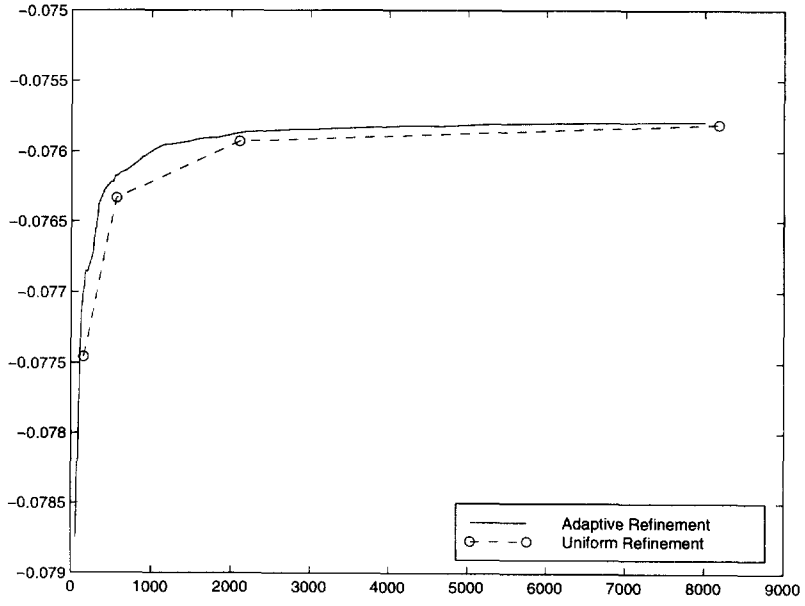


Fig. 21. $u_1(x_2)$ versus degrees of freedom in Example 8.3.

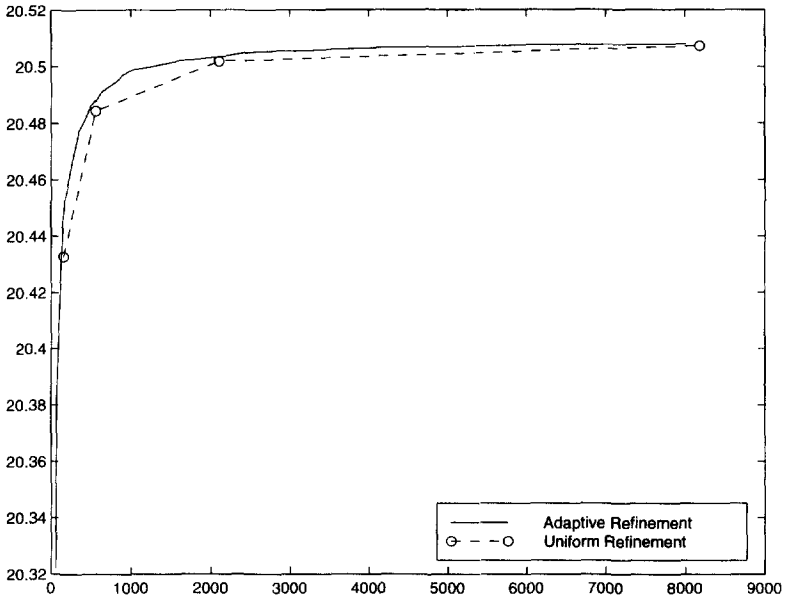


Fig. 22. $\int_{\Gamma_g} u_2(x)\,ds_x$ versus degrees of freedom in Example 8.3.

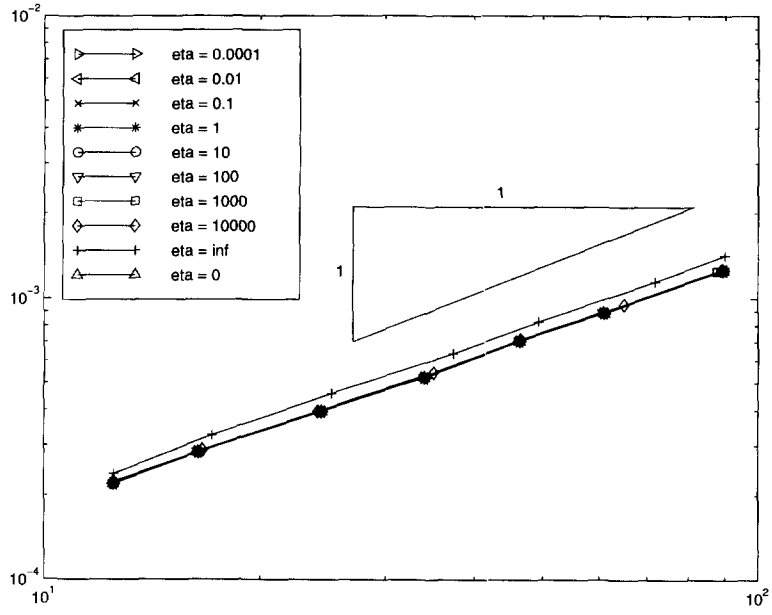


Fig. 23. $\eta(\mathcal{T})$ versus $\sqrt{n}$ in Example 8.3.

References

[1] W. Han and B.D. Reddy, Computational plasticity: the variational basis and numerical analysis, Comput. Mech. Adv. 2 (1995) 285–400.
[2] W. Han, Finite element analysis of a holonomic elastic-plastic problem, Numer. Math. 60 (1992) 493–508.
[3] C. Johnson, Existence theorems for plasticity problems, J. Math. Pures et Appl. 55 (1976) 431–444.
[4] C. Johnson, A mixed finite element method for plasticity problems with hardening, SIAM J. Numer. Anal. 14 (1977) 575–583.

[5] C. Carstensen, Numerical analysis of primal elastoplasticy with hardening Numer. Math., in press.
[6] G.A. Seregin, Differentiability of extremals of variational problems in the mechanics of perfect viscoplastic media, Diff. Eqns. 23 (1987) 1349–1358.
[7] G.A. Seregin, On the regularity of the minimizers of some variational problems of plasticity theory, St. Petersburg Math. J. 4 (1993) 989–1020.
[8] P.-M. Suquet, Discontinuities and plasticity, in: J.J. Moreau and P.D. Panagiotopoulos eds., Nonsmooth Mechanics and Applications, CISM Courses, Vol. 302 (Springer-Verlag New York, 1988) 279–341.
[9] R. Temam, Mathematical Problems in Plasticity (Gauthier-Villars, Paris, 1985).
[10] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, Introduction to adaptive methods for differential equations, Acta Numer. (1995) 105–158.
[11] C. Johnson and P. Hansbo, Adaptive finite element methods in computational mechanics, Comput. Methods Appl. Mech. Engrg. 101 (1992) 143–181.
[12] J. Alberty and C. Carstensen, Numerical analysis of time-depending primal elastoplasticy with hardening. Berichtsreihe des Mathematischen Seminars Kiel, Technical Report 98-18, Christian-Albrechts-Universitätzu Kiel, Kiel, 1998.
[13] C. Carstensen, Domain decomposition for a non-smooth convex minimisation problem and its application to plasticity, NLAA 4(3) (1997) 1–13.
[14] P.G. Ciarlet, The Finite Element Method for Elliptic Problems (North-Holland, Amsterdam, 1978).
[15] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods. Texts in Applied Mathematics, Vol. 15 (Springer, New York, 1994).
[16] C. Carstensen and R. Verfürth, Edge residuals dominate a posteriori error estimates for low order finite element methods. Berichtsreihe des Mathematischen Seminars Kiel, Technical report 97-6, Christian-Albrechts-Universität zu Kiel, Kiel, 1997.
[17] C. Carstensen, Weighted Clement-type interpolation and a posteriori analysis for FEM Berichtsreihe des Mathematischen Seminars Kiel 97-19 (1997).
[18] P. Clement, Approximation by finite element functions using local regularization, RAIRO Sér. Rouge Anal. Numér. R-2 (1975) 77–84.
[19] R. Verfürth, A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques (Wiley and Teubner, New York, Leipzig, 1996).
[20] J. Necas, Introduction to the Theory of Nonlinear Elliptic Equations. Teubner-Texte zur Mathematik, Vol. 52 (Teubner, Leipzig, 1983).
[21] J. Alberty, Ph. D. thesis in progress, University of Kiel, Germany, 1998.
[22] P. Wriggers et al., Benchmark Perforated Tension Strip. Communication in talk at ENUMATH Conference in Heidelberg, 1997. See also http://coulomb.mechanik.tu-darmstadt.de/user/scherf/Benchmark.html.
[23] C. Johnson, On plasticity with hardening, J. Math. Anal. Appl. 62 (1978) 325–336.